

The Evaluation of Students in the Classroom

A Handbook and Policy Guide

The Division of Evaluation
Department of Education
Government of Newfoundland and Labrador
1990
Authorized by The Minister





**THE EVALUATION OF STUDENTS
IN THE CLASSROOM**

A HANDBOOK AND POLICY GUIDE

**Department of Education
Division of Evaluation
St. John's, Newfoundland**

September, 1990



ACKNOWLEDGMENTS

The Department of Education expresses its sincere appreciation to members of the Student Evaluation Committee for their contribution to the development of this handbook and policy guide.

Judy Bartlett
Thomas Elementary
English Harbour West

William King
L.R. Ash Elementary
Lethbridge

Russ Blagdon (Chair)
Department of Education
St. John's

Dr. Mary Mulcahy
Memorial University of Newfoundland
St. John's

John Brisson
St. Peter's Academy
Benoit's Cove

Clarence Tucker
Gander Junior High
Gander

Jim Crewe
Pentecostal School Board
Windsor

Jim Dobson
Grand Falls Academy
Grand Falls

Jim Hill
Gander-Bonavista-Connaigre
R.C. School Board
Gander

Willis Jacobs
Department of Education
St. John's

Anna Kelly
St. John's R.C. School Board
St. John's



PREFACE

Student achievement is profoundly influenced by the evaluation practices used by teachers in the classroom. Research in education has consistently shown that achievement monitoring and feedback are extremely important factors in the instructional process and that they have a very significant impact on student success. Instruction can seldom be effective without a comprehensive evaluation plan that is carefully and systematically implemented in the classroom.

This handbook has been written to help provide classroom teachers, and others who provide direct support to classroom instruction, with information on good classroom evaluation practices. While it has a theoretical basis supported by references to evaluation literature, it is intended for practical use in policy development and in classroom application.

Evaluation is much more than testing. It is a continuous and comprehensive process, rather than sporadic and independent events. Evaluation should guide student learning daily rather than simply provide information for making decisions on promotion at the end of the year. Objectives are highlighted since they represent the beginning and the end of the instructional process, and there is a recognition that the process which students go through in coming to know is just as important as what they know. It is noted that comprehensive evaluation will very rarely occur when only one method of evaluation is used, and that if certain students are not to be at a disadvantage, special arrangements should be made in order to evaluate their achievement. It is pointed out that students should be graded in relation to the objectives that have been set for them, not merely in relation to other students. Two-way communications is emphasized since information must flow continuously between the home and the school, and it is noted that decisions on promotion or grade retention must always be based on a consideration of what is best for students, not only for the next year but for the rest of their lives.

The development of educational policy is an evolutionary process. This handbook, therefore, should be regarded as merely a starting point in the development of policy on student evaluation. The discussions of student evaluation can serve as a guide for the development of a district policy. The policy statements and guidelines which have been extracted from these discussions could serve as a guide for any district that wishes to modify, develop or implement a policy on student evaluation. As teachers work with it over the years, and educational researchers shed more light on complex educational issues, many of the ideas in this handbook will be refined, and the evolutionary process will have begun.



TABLE OF CONTENTS

Chapter I	Understanding Evaluation / 1
	• Meaning of Evaluation / 2 • Philosophical Bases of Evaluation / 3 • Principles of Evaluation / 3 • Summary / 4
Chapter II	The Evaluation Process / 5
	• Types of Evaluation / 6 • Objectives / 11 • Process and Product in Instruction and Evaluation / 18 • Differentiated Evaluation / 22 • Evaluation Issues in Effective Teaching Research / 26 • Summary / 31
Chapter III	Sources of Evaluation Data / 33
	• Homework / 33 • Learning Journals / 36 • Observation / 36 • Projects / 40 • Self-Evaluation Reports / 44 • Teacher-Student Conferences / 45 • Tests / 46 • Work Sample Files / 54 • Summary / 55
Chapter IV	Interpreting and Using Evaluation Data / 57
	• Grading / 57 • Reporting / 60 • Promotion / 63 • Summary / 65

Chapter V Special Issues in Student Evaluation / 67

- The Public Examination Process / 67
- Language Across the Curriculum / 68
- The Comparability of Grades / 69
- Time Spent on Testing / 69

Chapter VI Sample Policies / 71

- Evaluation of Student Achievement / 72
- Types of Evaluation / 73
- Objectives / 74
- Process and Product in Instruction and Evaluation / 76
- Differentiated Evaluation / 78
- Sources of Evaluation Data / 79
- Grading / 82
- Reporting / 84
- Promotion / 86

Reference List / 87

- Appendixes:**
- Levels of the cognitive domain with appropriate verbs / 97
 - Levels of the affective domain with appropriate verbs / 99
 - Levels of the psychomotor domain with appropriate verbs / 100
 - A table of specifications / 102
 - A general evaluation plan / 103
 - Checklist / 106
 - Anecdotal Record / 107
 - Rating Scale / 108
 - Project Assignment Sheet / 109
 - Student Self-Evaluation Form / 110
 - Constructed-Response Items / 111
 - Selected-Response Items / 106
 - Low-order and high-order selected response items / 117
 - Grade retention scale / 121

CHAPTER I

UNDERSTANDING EVALUATION

Evaluation of student achievement is a crucial and pervasive issue which is of concern in every classroom. Traditionally, evaluation activities were used to certify or promote students. Currently, evaluation is recognized as a comprehensive process, not just something that occurs only at the end of some learning episode. It is an integral part of the teaching-learning process. Evaluation activities provide information for the identification of students' strengths and weaknesses, a basis for the assessment of teaching strategies and data for important decisions regarding corrective or enrichment activities.

All teachers are concerned about how well students are achieving course objectives, and how much they influence students' cognitive, affective and psychomotor development. They realize that it is possible to meet students day after day and have little substantive influence, but it is also possible to affect profoundly their future development. If the classroom is to function effectively, a well-planned system of evaluation is essential. Without such a plan, it is difficult to determine the abilities of students, and to ascertain the methods of instruction most appropriate for a particular class or group of students. It can have serious consequences for the teacher, even though the effects are not personal and only temporary. For the student however, the effects of the evaluation may last a lifetime. This is

pointed out by Napier (1976), when he stated:

Do you consider yourself fairly smart or not? And when did you perceive yourself in this way? Chances are that your vital self-concept germinated early in your school career. Possibly it was the onset of your tortuous grading history. It is likely that some teacher was simply performing his or her duty by giving you a grade of some sort. He or she did not intend to assign you a life-long membership to the smart, the dumb, or the just plain old average group. But there you are twenty years later. (p.23)

Determining the extent to which students are achieving the educational aims and objectives is a difficult process, but it is central to teaching. Proper classroom evaluation can improve overall classroom achievement, the students' self-confidence, teacher-student relations, and teaching. On the other hand, improperly managed evaluation can stifle learning and retard both student and teacher growth.

Although students evaluation is, to some degree, based on intuitive feelings about where students stand in relation to course objectives, intuition alone is not enough. Evaluation based primarily on intuition can be capricious and superficial. One of the products of such evaluation is diff-

iculty in communicating results. Also, many details and nuances tend to be overlooked, and decisions based exclusively on intuitive evaluation may not always be the best for a particular student at a particular time. Thus, while feelings and intuition are important in student evaluation, a systematic and purposeful program is essential.

It is stated in the Quebec General Policy for Educational Evaluation that "when one evaluates a pupil's learning or general development, one makes a judgement about a human being who has an inalienable right that such judgements be as soundly based on reality as possible" (Quebec Ministry of Education, 1981, p.4). It should be remembered that evaluation decisions affect more than teaching methodology and course objectives. Students are an integral part of the process, and they have a right to expect that decisions made about them are based on the best possible information.

In this chapter, the differences between evaluation and measurement are noted and the philosophical foundation for a comprehensive and effective evaluation policy is provided. While many principles and beliefs are implicit in the philosophy, more precise and explicit statements that should guide the evaluation process are provided.

Meaning of Evaluation

The term *evaluation* is frequently interchanged with the terms *measurement* and *testing*. These terms have distinct

meanings. Evaluation is the systematic process of collecting, analyzing and interpreting data; measurement is a numerical description of an individual's performance; and testing is only one possible instrument that can be used to collect data. Gronlund (1985) stated "evaluation is a much more comprehensive and inclusive term than measurement is, and testing is just one type of measurement" (p.5). He pointed out that evaluation may involve quantitative descriptions (measurement) and qualitative descriptions (nonmeasurement) of pupils, where the main emphasis is "on the extent to which learning outcomes are achieved" (p.6).

Measurement seeks to answer the question "how much?" Evaluation is a much more comprehensive process which seeks to answer the question "of what value?". A good evaluation system relies on more than measurements. This is not to suggest that quantitative information



relating to student achievement is bad and that qualitative information is good. Rather, both are necessary for comprehensive student evaluation.

Evaluation is, essentially, a process involving a value judgement regarding the achievement of stated objectives. It is a comprehensive process encompassing the entire range of students' activities and experiences. It is continuous in that there is no starting, no finishing point. There are no blank periods in evaluation just as there are no blank periods in human development. The evaluation process takes place at the beginning of a unit of work, during the course of the unit, and at the end of the unit. It is based on the assumption that appropriate tools of measurement will be used whenever needed.

The fundamental question that any adequate theory of evaluation must address is not what can be evaluated, or how or whether objectives have been achieved, but how it is that human beings come to know and, how it is that they communicate what they know to others. (Willis, 1978). This would seem to imply that we need to know about the nature of human beings and the nature of knowing.

Philosophical Bases of Evaluation

Education is concerned with the integral formation of students. Hence, it should be directed toward their physical, emotional, aesthetic, social, moral, intellectual and spiritual development.

Objectives designed to guide evaluation should take direction from this multidimensional development. An understanding of the nature and truth value of knowledge is fundamental to an understanding of the learning process and, consequently, of the evaluation process. Any concept of human knowing takes in the whole of human life: the external senses, the imagination, the feelings and emotions, the ability to think and reason and to make responsible decisions. This has implications for teaching and learning and for evaluation. Objectives should be grounded in a theory of knowing which recognizes the multidimensional nature of human beings. There are objectives which deal with facts, critical thinking, values, attitudes, appreciations, and beliefs. There are objectives that are concerned with study skills, group skills, and the many manipulative and manual skills that form part of the school curriculum. It follows, then, that any system of evaluation has to be based on who students are and how they know and act in their pursuit of the true, the good, and the beautiful.

Principles of Evaluation

The following general principles of evaluation should be evident from the philosophy and are more fully developed in various sections of the document.

- Evaluation takes its direction from a definition of education which, stated in broadest sense, is to enable students to realize their potential as human beings.

- Evaluation is a means to an end and never an end in itself.
- Evaluation, by definition, connotes value.
- The process of evaluation begins with the objectives of the educational program.
- The primary purpose of evaluation is to determine to what extent students have achieved the objectives of the educational program.
- Evaluation gives a picture of where students are and how they are progressing.
- Evaluation has some attributes which are quantitative and some which are qualitative.
- Evaluation is a more inclusive concept than measurement - measurement provides data needed for evaluation, but it is not evaluation per se.
- Measurement implies the use of some tool or instrument such as a test or scale and provides a quantitative description of observed phenomena.
- Instruments used for evaluation should meet the criteria of validity and reliability.
- Evaluation is a continuous and comprehensive process.
- Every evaluation should be made with reference to specified criteria of some kind.
- Evaluation procedures should take into consideration individual differences among students.
- Evaluation procedures are important for the individual guidance of students.
- Evaluation of students involves more than a single appraisal at any one time.

Summary

Student evaluation is a process basic to teaching. Evaluation is not an add-on feature of instruction but an integral part of it, since the information it provides allows teachers to make adjustments to objectives and teaching methodologies.

Evaluation must be student-centered where the process and subsequent decisions reflect a genuine concern for the interest of each student. This can be realized by adhering to the basic philosophy which has been outlined, and by recognizing that evaluation is designed to help a student.

Teachers must recognize that they have a responsibility to provide an accurate, reliable and justifiable evaluation, which does not undermine the integrity, self-confidence, or enthusiasm of the student. Comprehensive evaluation involves the use of a variety of sources, with reference to specific and rational criteria. It must be implemented in a manner which ensures that the total growth of every student is recognized.

CHAPTER II

THE EVALUATION PROCESS

Evaluation of student progress does not take place in a vacuum. It takes place in the context of teaching and learning, and many factors in this context have implications for student evaluation. As Figure 1 illustrates, the formulation of objectives, the selection of teaching strategies, and evaluation are ongoing activities which may be repeated throughout the teaching-evaluation process.

Formative evaluation activities are used to identify impediments to learning and to ensure that learning is taking place. Summative evaluation is used in certifying students' levels of achievement.

Instruction and evaluation are centered around objectives. Not only are objectives used in providing structure for teaching and learning, but they also provide a framework for evaluation. In most cases,

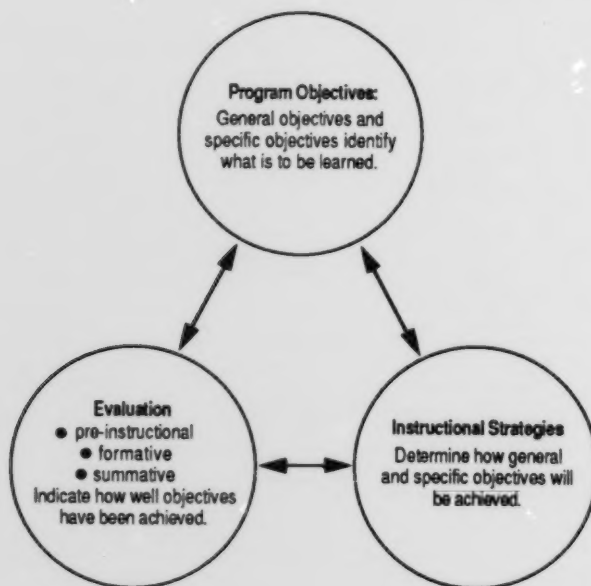


Figure 2: The Teaching-Evaluation Process

Evaluation of student progress can be pre-instructional, formative or summative. Pre-instructional evaluation is designed to ensure that instruction is appropriate.

objectives specify a product. However, a learning environment where the emphasis is on product only is likely to be sterile and unchallenging for many. In any

classroom, there should be a balanced approach to instruction in which consideration is given not only to the products of learning, but also to the learning process.

An education system must meet the needs of all students. Just as curriculum must be adapted to meet special needs, evaluation must be adaptable to special needs. Otherwise, a system will be created where some students cannot hope to survive, and others will not be challenged.

Researchers have recently shed much light on the relation between student achievement and characteristics of teachers and schools. Those research findings that relate directly to evaluation, along with types of evaluation, objectives, process and product, and differentiated evaluation are discussed in this chapter.

Types of Evaluation

For the purpose of this discussion, three types of evaluation are defined: pre-instructional, formative, and summative. Each of these have different purposes. To the casual observer, one type of evaluation looks much like another because the type of evaluation is determined by how the results are used. If results are used to determine appropriate starting points for instruction, evaluation may be regarded as pre-instructional. When data are used to determine how well instruction is going, it

is formative evaluation. If the purpose is to determine the extent to which students have achieved the objectives of a unit of study, evaluation is summative. However, where it is decided after a summative evaluation that students have not achieved objectives to a satisfactory degree, and a decision is made to reteach, evaluation intended as summative becomes formative.

Pre-instructional Evaluation

In the pre-instructional phase of evaluation, an attempt is made to determine what each student brings to the learning situation or where each student is in relation to stated objectives. This is the essential first step in curriculum adaptation and the proper placement of students. As Gronlund (in Shapiro, Lukesavith & Shapiro, 1986) indicated, the better the identification of needs, the more direction there is for instructional and curricula change or both. However, caution must be exercised to ensure that pre-instructional evaluation does not have a negative impact on student performance. Pre-instructional evaluation must be seen as an indicator of where students are, and not as a measure of what they are capable of doing. Schuncke (1981) noted that errors occurred when pre-instructional evaluation was done in a very informal and intuitive manner. It frequently resulted in teaching that was predicated upon a number of false assumptions about a student's readiness or ability to learn. The result was frustration and failure for both student and teacher.

Closely associated with the pre-instructional phase of evaluation is the issue of expectations. A number of studies (Brookover et al. 1982; Good & Brophy, 1978; Rosenthal & Jacobson, 1968) have examined the effects of prior expectations, and have concluded that prior expectations may affect performance. However, caution must be exercised. The purpose of pre-instructional evaluation is to obtain information to guide the formulation of instructional strategies which may enhance the performance of students. It is not used to create negative expectations, since teachers should avoid holding prior expectations which may be detrimental to a student's performance. Teachers should believe that adjusting instructional strategies will enhance the performance level of students, and no matter where students are on the achievement continuum, reasonable and appropriate educational experiences should be provided.

If those students who are expected to have difficulty do not get an equal opportunity to take part in classroom activities, they are likely to fall further behind their classmates as time passes. A vicious circle develops in this situation. As students fall further and further behind, teachers expect less and less, and as teachers expect less and less, students fall further and further behind. What, then, is a teacher to do? They may desire to hold no expectations for any students, or hold high expectations for all. It is difficult to hold no expectations because even if teachers avoid cumulative records and

comments from previous teachers, they will begin to build expectations as they interact with students. Furthermore, teachers need information on past experiences of students in order to plan appropriately. While a positive attitude can make a difficult task easier, the same expectations should not be held for everyone. Some students are capable of achieving at a higher level than others, and holding the same expectations for everyone will result in some students experiencing frustration. Furthermore, it is a distortion of reality to assume that everyone can achieve at the same level. What is important is that expectations be appropriate. An appropriate expectation for all students would be one that sees them working to their potential and progressing at a steady rate.

The basic goal of instruction is always to promote learning and development in the most effective and efficient manner possible. Each program, course, or lesson begins at a distinct point determined by the calendar or the curriculum. To avoid frustrating some students with unreasonably high expectations, while boring others with unreasonably low ones, it is necessary to determine where students are in relation to the planned objectives. As the work of Piaget suggests, the most skilful presentation of a concept is likely to fail if students are not yet ready to assimilate the concept.

Pre-instructional evaluation serves a number of purposes. It provides information on the level of entry skills and knowledge

of each student in relation to objectives and it may also indicate the learning styles of students. As well, it may indicate the need for particular modifications of the instructional program. Teachers must be fully aware of the knowledge base, plan an appropriate teaching and evaluation process, and subsequently modify these plans as the process unfolds.

There are a number of sources available for pre-instructional assessment of entry skills, knowledge, attitudes, and behaviours. These include cumulative records, standardized and teacher-made diagnostic and achievement tests, meetings with parents and teachers, and teacher-student conferences. Teachers should use a variety of these sources to acquire the necessary pre-instructional information, continuously updating the information during the teaching - learning process. Even if students do not know something today, it does not mean that they will not know it tomorrow.

Formative Evaluation

Formative evaluation is the evaluation of "how things are going" rather than a decision about "how things have gone". It is evaluation that occurs *during* the teaching - learning process. Andrews, Gardner & Hubbard (1986) defined it succinctly as "the procedure by which a teacher structures interactions with students in order to obtain specific information regarding progress in the

...course material" (p.8).

Formative evaluation is concerned more with process than the product of learning. Its fundamental purpose is the identification of a student's strengths or weaknesses with respect to specific objectives so that necessary modifications can be made. The diagnostic aspect is dominant in formative evaluation where emphasis is on identifying and eliminating barriers to success.

A systematic process for monitoring student progress is required in formative evaluation. The value of immediate and continuous feedback to the learner is well-documented in the research on effective teaching. However, specialist intervention and remedial assistance should be provided for the learner when expert knowledge is needed. Such interventions, when used appropriately in formative evaluation, can be of immediate benefit to the student (Lavi, 1982).

Although communication is a key element of formative evaluation, it is important that teachers share with colleagues their particular knowledge and insights with respect to the needs of specific students. It is also essential that parents be well-informed about the progress of instruction. There are few parents who do not want what is best for their children since all parents have a large stake in their children's success. They have a right to know how their children *are doing* as opposed to how their children *have done*,

so that they may take an active role in providing the best possible learning situation.

An underlying assumption of formative evaluation is that in learning, the responsibility for success should be shared. Principals, teachers, parents, and students generally agree on the purposes of schooling, and all are dedicated to the proposition that children can and will learn. Parents should not assume that their children's education is the complete responsibility of the teacher. Children have to assume some responsibility for their own learning, and parents have an obligation to encourage them to do well.

It is widely recognized that in today's information age when the sum total of knowledge is, at a very conservative estimate, doubling every ten years, education must concern itself more with the process of learning than with its specific outcomes. Therefore, the teaching methodology should move away from the chalk-and-talk method of dispensing information. An interactive orientation should be emphasized wherein the student learns by doing. In this methodology, the emphasis should be placed upon developing the capacity of the learner to acquire and use new knowledge. The facts are the means to an end; they are not ends in themselves. If such a methodology is to be implemented successfully, the necessary formative evaluation data should be obtained through extensive use of a variety of sources.

In a study of the effects of formative evaluation in classrooms, Mirkin and David (1979) concluded that "systematic

formative evaluation most effectively contributes to student achievement when rules for the utilization of measurement data are included as part of the formative evaluation system" (p.12). Measurement, through the use of tests, is an important part of evaluation. However, there are other equally useful and often less time-consuming methods of assessment that should be used in formative evaluation (observation, work sample files, and so on). Overemphasis on testing as an evaluative tool should be avoided.

Formative evaluation may be conducted by the teacher, or by students themselves as they evaluate their own, or their classmates' work. Teacher-conducted formative evaluation provides information on the achievement of objectives. It indicates the strengths and weaknesses of students, program, and teaching methodology. As well, it provides work samples that might be useful as part of the reporting process. When students evaluate themselves, they can develop a sense of responsibility for their own learning. They become aware of objectives and develop the capacity to evaluate their own work. As well, this type of evaluation can provide feedback to the teacher on student reactions to various aspects of the course. Peer evaluation provides students with responses to their work other than the teacher's, and it gives students an opportunity to compare their work with that of other students.

Summative Evaluation

Summative evaluation is used to grade, certify, or select students. It is not

intended to improve current instruction for the benefit of those being evaluated. It measures what has happened, not what is happening. But results can serve to indicate areas of strength or weakness in subjects involving sequential skill development. These results may be used to influence later instruction. In that sense, summative evaluation will have a formative dimension. For evaluations having a summative component, students should be informed in advance of the event as to its timing, method, and the objectives which are to be evaluated.

Summative evaluation is the evaluation of student achievement at the end of instruction. It is directed towards a general assessment of the degree to which outcomes are attained over the entire course, or some substantial part of it. The essential characteristic is that a judgement is made about a student with regard to what learning has taken place or is supposed to have taken place. Summative evaluation may be used for diagnosis as well as for deciding who is permitted to go to the next level. It may be used to determine the present status of students, to identify factors which are responsible for and influence growth and development, and to determine potential for future growth and development.

Data for use in summative evaluation can be obtained from a variety of sources. At the primary and elementary level, teacher observation is a major source, while testing is used more frequently at the intermediate and senior high levels. In any event, the primary purpose is to gather sufficient data to ensure that a valid judgement is made about a student's

achievement of stated objectives. For example, many French courses and the research course in the senior high school program cannot be appropriately evaluated using paper-and-pencil tests. Sources of evaluation data such as work samples, interviews, and observations may also function as summative instruments, if used separately for this purpose. Instruments used in summative evaluation should reflect the methods of instruction and the significance and sequence of the objectives. The student's evaluation should be referenced to the objectives of instruction rather than simply to relative rank among those evaluated. The instruments used should be flexible enough to permit a variety of student responses, and should provide scope for the demonstration of divergent thinking consistent with the objectives of instruction.

The tendency to average formative results according to a rigid, predetermined formula for summative purposes should be avoided. Sadler (1983) noted that the practice of 'counting' everything a student submits towards a course grade improves the sampling but, too often, at the cost of improvement. In other words, like the practice of averaging student results on tests, it can result in an unrealistically low appraisal of the student's knowledge in relation to the objectives of instruction. Also, this practice penalizes the learner who thoroughly achieves the objective at a late stage of the instructional program. It is essential that students have the freedom to err and profit from mistakes without fear of penalty or reprisal.

The best person to evaluate the progress of a student is the person responsible for

instruction. It is the summative phase of evaluation which calls for the most careful exercise of the teacher's professional judgement. As stated earlier, evaluation is not synonymous with measurement. Evaluation information must be sensitively interpreted by an informed evaluator if the right decision is to be made.

Objectives

Objectives are goals defined in terms of human thoughts and performances. These objectives spell out actions and values that are deemed to contribute to the individual and to society. Of paramount importance is respect for the worth and integrity of all students with due concern that all be given an equal opportunity to grow and develop to the fullest extent of their capabilities. Thus, a logical step in discussing evaluation is to examine the role of objectives in guiding the teaching-learning process.

Education is a cumulative process in that everything a student is taught or is expected to learn is geared towards the total development of the individual. Just as a teacher should have a vision of where a student should be at the end of a sequence of instruction, an education system should be supported by a vision of the type of individual who is truly educated. In this province, this vision is articulated in the *Aims of Public Education for Newfoundland and Labrador*. However, as Zais (1976) indicated, aims are outcomes that are removed from the

school situation to such an extent that their achievement cannot be determined until well after schooling is finished. The educational aims give rise to a specification of more narrow outcomes of schooling. These are often referred to as goals. These goals specify the general outcomes of schooling and are written as general objectives in the *Aims of Public Education for Newfoundland and Labrador*.

One vehicle which helps move an education system toward achievement of aims and goals is curriculum. Curriculum is chosen or developed so that achievement is affected most positively. Outcomes which are achieved over a significant portion or all of a student's school life are often referred to as *curriculum objectives*. For example, a general curriculum objective for social studies is students should understand how geographic conditions affect people's way of life. This general curriculum objective is achieved through shorter-range *course objectives* emanating from the Grade 4 study of life in selected world communities. The course objectives, in turn, are achieved through still shorter-range *lesson objectives* relating to a study of life in a given community. Lesson objectives on winter activities in that community lend specificity to the major understanding stated above.

Since curriculum is centralized, general curriculum objectives are specified by the Department of Education. The objectives are provided through provincially-adopted curriculum guides. At the level of the lesson, however, teachers have to form-

ulate their own objectives or rely on those provided by educational publishers. It is those objectives that focus the attention of students on relevant and important material to be learned, and the attention of teachers upon materials and procedures used to achieve them. Thus, objectives guide both teaching and learning, and ultimately they guide evaluation, since the purpose of evaluation is, after all, to determine the degree to which objectives have been achieved.

If objectives are to guide learning, students should be aware of them. They are most meaningful if they are written in language that students understand and if they specify what students should be able to do as a result of instruction. These two criteria would be useful in analyzing objectives to determine their suitability for use in the classroom.

Objectives may be very simple, or they may be very complex. They may be achievable or unachievable. They may be stated explicitly, presented in great detail, or they may be inherent in the material or experience or in thought processes, values and judgements. (Dressel, 1968)

School programs specify many kinds of objectives which students are expected to achieve. There are objectives that specify intellectual development, those which specify the development of values, feelings and attitudes, and those that specify the physical development of the individual. Bloom (1956) developed a taxonomy which classifies all objectives into three domains: cognitive, affective, and

psychomotor. In order that the total development of students be emphasized, it is important that all three domains be addressed when instruction is planned.

Eisner (1985) identified instructional objectives and expressive objectives. Instructional objectives specify a particular outcome, skill, or item of knowledge that the student should acquire on completion of one or more learning activities. Expressive objectives describe an educational encounter. They identify a situation in which students work with a problem or engage in a task, but it is not specified what they are to learn from that encounter, situation, problem, or task.

Expressive objectives provide students with an invitation to explore or focus on issues of individual interest. They are evocative rather than prescriptive. Wolf (1979) gave three major sources of objectives: the needs of students, the needs of society, and the material to be learned. All three are integrally connected.

Objectives have some limitations. Among others, the following may be noted:

- Prediction of educational outcomes cannot be made with accuracy.
- Subject matter affects precision in stating educational objectives.
- Some educational objectives cannot be measured. They provide only a criterion for teacher judgement.

Despite these limitations, however, it still

remains that objectives provide the rationale for the whole student evaluation process.

While it is generally accepted that objectives and goals are essential to effective instruction, there is not always agreement on how objectives should be written. Stated as an instructional intent, an objective might read, "To demonstrate to students how to construct a model of the solar system". The difficulty with this kind of statement is that it does not specify what is expected of the student and, literally speaking, the objective is achieved once the demonstration is completed. A more effective way to state the objective would be, "Following a demonstration, the student will be able to construct a model of the solar system showing the relative size of the planets and their distances from the sun". This objective focuses on students and what they should be able to do following instruction.

This distinction between stating objectives in terms of teacher intent or in terms of learning outcomes is an important one. Although it is recognized that there is some value in having teachers know their instructional intent, it must be viewed as a preliminary step to translating that instructional intent into objectives which focus on learning outcomes. The whole purpose of instruction is to bring about changes which indicate that learning is occurring. In this context, evaluation of the objectives determines the degree of achievement of learning outcomes.

Gronlund (1970) identified the following as ways in which objectives, stated as learning outcomes, contribute to the learning process:

- They provide a guide for student learning activities both in and out of the classroom.
- They provide direction to teachers and clearly convey their instructional intent to others.
- They provide a guide for selecting the subject matter, teaching methods, and materials to be used in instruction.

If the first step in instructional planning is the identification of objectives that support long-term goals, then the second step is how to write or select those objectives. These objectives must identify intended learning outcomes without being either so specific that they become ends in themselves or so general that their vagueness fails to achieve the intended purpose. Objectives which are too specific have some serious limitations. They may become ends in themselves, and there is a danger that the emphasis will be on recall of knowledge. This should not be the focus of instruction. The function of teaching, and hence the function of learning, is not to provide students with facts for recall upon demand. Rather the purpose of education, as stated by John Dewey, is to allow individuals to come into full possession of their personal power. This can be achieved only if teachers have a clear vision of what

constitutes an educated person, and if they design instruction to help realize this vision.

Gronlund (1970) proposed a two-step process for writing objectives.

- State the objective as a learning outcome. For example, "students will understand the meaning of literary terms".
- List under each objective a sample of the specific learning outcomes that indicate the attainment of the objective. For example, following from the above objective, students should be able to define the term in their own words, identify the meaning of the term when used in context, and distinguish between terms having similar meaning.

Emphasis should be on broad objectives, since they identify the quality of learning outcomes towards which instruction should be geared. At the same time, there is a recognition that specific indicators are the measure of the attainment of those objectives. Clearly stated objectives should ensure that instruction goes beyond the simple recall of information so that the larger domain of learning is addressed.

The remainder of this section deals with the writing of objectives in the cognitive, affective, and psychomotor domains. This is not to suggest, however, that learning in any one of these domains is independent of learning in another. Learning generally

involves all three domains simultaneously. They are organized as separate domains here for purposes of the discussion only.

Evaluating the Cognitive Domain

The cognitive domain emphasizes the acquisition of skills that focus on knowledge, understanding and thinking. Bloom, in his taxonomy, identified six major categories of learning in the cognitive domain. Learners recall previously learned material; they understand, apply this knowledge and analyze it. They put parts of it together to form a new whole, and they judge the value of material for a given purpose. These categories are hierarchical and represent increasingly more complex forms of learning. The taxonomy is also perceived as encompassing all possible learning outcomes that might result from instruction.

A distinct advantage in being able to classify objectives through the use of a taxonomy is to ensure that the focus of instruction goes beyond the lower levels of learning. While all levels may not be covered for each segment of instruction, a knowledge of a taxonomy will help prevent serious omissions while reminding the teacher of the complexity of learning outcomes to be considered. Teachers must ensure that the full cognitive abilities of students are developed. To assist teachers in writing objectives and evaluating higher-level cognitive skills, a list of appropriate verbs for the six levels is presented in **Appendix A**.

The gathering of evaluation data about the cognitive domain is fairly straight forward, because objectives can be expressed in terms which make them observable. Practically any source of evaluation data can be used to determine student performance in the cognitive area. However, teachers must ensure that evaluation goes beyond the recall of knowledge.

Evaluating the Affective Domain

The affective domain deals with feelings and emotions and is exemplified through students' attitudes, interests, value systems, and appreciations. There is little doubt that this is a more nebulous area than either the cognitive or psychomotor domains. In the affective domain, emphasis is on the degree to which students develop age-appropriate and socially acceptable behaviour. For this reason, it is neither appropriate nor desirable to attach numerical values to affective outcomes. Furthermore, students' progress in the affective domain should not influence their progress in the other domains. This does not, however, diminish the importance of the affective domain. As stated by Eiss and Harbeck (1969):

The affective domain is central to every part of the learning and evaluation process...It includes values and value systems that provide the basis for continued learning and for most of an individual's overt behaviours. It

provides the bridge between the stimulus and the cognitive and the psychomotor aspects of an individual's personality. (p. 11)

The immediate question is, why has so little attention been given to assessing student progress in such a crucial area of child development? There may be a number of reasons: perhaps it is expected that the home and church will continue to provide instruction in those values; perhaps it is because schools tend to place emphasis on the cognitive aspects of education, and perhaps the need for affective development is too often overlooked in statements of course objectives. Undoubtedly, a major reason is the difficulty associated with measuring progress toward affective goals. Teachers are somewhat reluctant to deal with delicate, personal, and controversial values. Whatever the reasons, it is obvious that very little attention is given to the development and evaluation of objectives in the affective domain. While teachers are concerned that students develop an "appreciation" for poetry, instruction almost always focuses on cognitive aspects, resulting all too often, in an adverse affect - an aversion for rather than an appreciation of poetry.

Just as the classification of objectives in the cognitive domain helps focus on all aspects of intellectual development, a similar classification in the affective domain should be of invaluable assistance. Krathwohl (1964) identified five categories for classifying affective objectives.

Learners pay attention to a stimulus, respond to it, and place a value or worth upon it. Values are brought together and a value system that is internally consistent begins to be developed. Behavior which is consistent and predictable is governed by this value system. Even with such a classification, it is still somewhat difficult to evaluate outcomes in this domain. Eiss and Harbeck (1969) stated that this difficulty exists because a given behaviour may indicate the attainment of any one of several objectives depending upon the thinking and motives of the individual exhibiting the behaviour. When students produce the desired behaviour on demand, it often results in the behaviour itself becoming the goal rather than an indicator that the goal has been attained.



The announcement of specific objectives in advance poses a problem since it cues

students to display particular behaviors. This does not mean that affective objectives are to be treated in an incidental manner or neglected altogether. On the contrary, they become even more important. For, while we might attempt to observe students in unstructured situations to determine the presence of desired responses, we must also provide students with opportunities to develop those qualities deemed acceptable as social norms. This can be accomplished only through a deliberate attempt to develop appropriate objectives and use means other than traditional instruments (i.e. pencil and paper tests) to evaluate outcomes. It is not acceptable to wait and hope for students to exhibit desirable affective behaviours. There must be a deliberate attempt by teachers to foster appropriate behaviours.

To assist teachers in writing affective objectives a list of appropriate verbs for the five categories is presented in **Appendix B**. When writing these objectives for the affective domain, the teacher must be cognizant of the need to define expected outcomes as clearly as possible. In this regard, the two-step process outlined for the cognitive domain can be extremely helpful.

Perhaps the most effective method of evaluating objectives in the affective domain is observation. This method, utilizing appropriate recording devices such as rating scales, checklists, and anecdotal records, enables the teacher to focus on and record performance. Other sources of evaluation data for the affective domain

include self-reporting devices such as an attitude scale where students make responses on a continuum ranging from favorable to unfavorable, and individual inventories where students check their preference for or attitude towards an idea or activity. Student reports or term papers can be evaluated from an affective point of view if they are designed so that the student is required to take a stand on a certain issue. Personal interviews and discussions can also be used to evaluate affective growth if care is taken to use indirect questions so that students cannot guess which answer the teacher would like to hear.

Evaluating the Psychomotor Domain

The psychomotor domain focuses on objectives that emphasize motor and manipulative skill development. Although this domain may contain objectives which are common to most subjects, (such as writing in language arts, construction in mathematics, using laboratory equipment in science, painting or modelling in art, and map drawing in social studies), the emphasis is on performance skills and most attention is given to it in subjects like home economics, industrial arts, physical education, and music.

The development of a classification system from the psychomotor domain is more recent than those for the cognitive and affective domains. Simpson (1972) proposed a classification scheme containing seven major hierarchical categories. The learner uses sense organs to obtain cues, displays readiness to take a particular action, deals with a complex

skill by imitation as well as trial and error, habitually responds to a performance act where movements are performed with confidence and proficiency, performs complex movements with skill, modifies movement patterns to meet special requirements or problems situations, and creates new movements to fit a particular situation or specific pattern. A further description, along with sample objectives and specific learning outcomes is provided in **Appendix C**.

The rationale for using observation as the most effective source for evaluation data in the affective domain applies equally to the psychomotor domain, although some measurement may be used in subject areas such as physical education. Once teachers have identified desirable outcomes, careful observation using appropriate recording devices can be used effectively to evaluate them.



Process and Product in Instruction and Evaluation

Debate about the relative importance of process and product in instruction and evaluation is quite common. The controversy implies that there is a dichotomy between the two orientations, and quite often one's experiences reinforce the perception that a gap does indeed exist. On the one hand, it is perceived that education consists of having students "stockpile" knowledge deemed to be of most worth. On the other hand, it is argued that greater attention should be given to the kinds of processes most conducive to the acquisition and utilization of knowledge. An attempt will be made here to define the product and process orientations to instruction, and to examine possible implications for the evaluation process.

Understanding Process and Product

Product can be thought of as the *what* of learning. *What* is it that students should know? *What* is it that they should be able to do? These *whats* are typically spelled out in objectives, whether they are based on a program, a course, a unit or a lesson. In a teaching episode which is biased towards product, the ends are evaluated without reference to the means. All is well as long as students know or can do what the objectives say they should know or be able to do. Generally speaking, students are passive in teaching situations oriented toward product.

Process refers to the procedure which students go through as they come to know, to value, and to be able to do. The emphasis in teaching shifts from a passing on of "*the what*" to the methods by which it is acquired and utilized. Students are taught to find information, to become autonomous thinkers, and to use knowledge to solve new problems and make decisions for themselves. In a process-oriented approach to instruction, students are active participants.

The differences between process and product-oriented approaches to instruction are perhaps better understood by looking at an extreme example of each. In a social studies course, there is a body of content which students are expected to know. The teacher provides students with what is thought to be an exhaustive list of questions, along with the answer to each question. Students are required to "learn" this and reproduce answers to a sample of questions appearing on an examination that is given at the end of some episode. In this situation, the emphasis is solely on product. The process that many students will go through is memorization. Any reasonable means of producing the product is justified, although actual learning experiences may be minimal.

A corrective reading program used in some schools many years ago was an example of a program which was heavily oriented toward process. Students were taught decoding skills and required to use rules to determine such things as whether a word had a long or a short vowel sound. Learning was supposedly demonstrated

when a student could read nonsense words. Students were learning to read, but the program provided nothing to read other than nonsense words. In this case, a product (or content) was completely absent and learning was likely minimal as well.

Frost (1989) stated, "...quality programming and instruction are neither content based nor process based, but a wise and judicious mixture of both" (p.11). On the one hand, it is important that students learn to produce a product; on the other, how they learn to produce that product is equally important. The dichotomy is more the degree of emphasis given one over the other. As Figure 2 shows, a teacher at point A on the process-product continuum would view product as most important. Process would be given almost no attention. A teacher at point C would view process as being so

important that a knowledge base essential to full development may be lacking. A teacher at point B would have struck a balance between process and product.

There should be congruence between what is taught, how it is taught, and the emphasis in the evaluation process. Let's suppose, for example, that the oral component of a French program receives more emphasis than the written. If the evaluation at the end of the semester gives a relative weight of 50-50 to the components, then it reflects a lack of congruence. According to Pratt (1980), students, unless they have been informed, have a right to feel that they were misled as to what would be evaluated. Teachers need to ensure a balanced representation of all major topics and consider the emphasis given to product and process. This can be accomplished by developing a table of specifications (see Appendix D). Such a table can also ensure that the various levels of cognitive skills, indicated by stated objectives, receive the appropriate emphasis. Product evaluation should not be confined to the recall and comprehension levels, but should include the processes of analysis, synthesis and evaluation. However, at the higher levels, process gains more significance if students are to use the knowledge which they have acquired, rather than simply recall it.

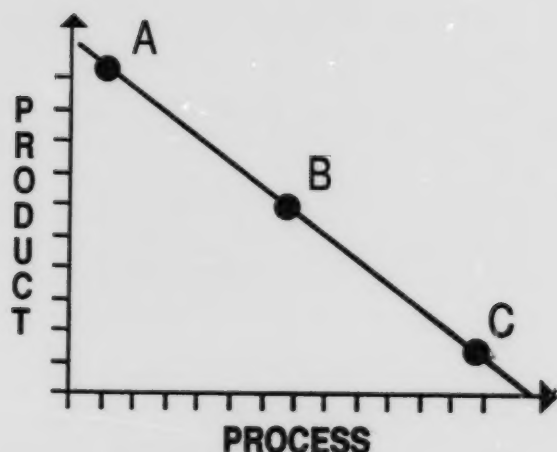


Figure 2: The Product-Process Continuum

Evaluation of Process and Product

There are some interesting questions which arise regarding the evaluation of process and product. Both should be evaluated, but they cannot be evaluated independently, and the evaluation of each implies a different use of the available

data. Conclusions regarding the efficacy of a process can usually be reached only through the evaluation of product. If the product is good, the process is usually acceptable. If the product is unacceptable, the efficacy of the process is questionable.

While the evaluation of process is usually formative, it can be summative when the process is the objective. For example, it might be an objective of a social studies course that students experience living in a model democratic environment. If such an environment is created, and students experience living in it, the objective is achieved. When the process becomes the product, teachers need only observe that a student has been involved in the process. A satisfactory evaluation can be provided without any reference to the quality of that experience. If, however, students are to learn about such specific concepts as majority rule, there will have to be an evaluation of product to determine if students have learned what they are supposed to have learned. This does not imply that the evaluation has to be of the pencil-and-paper type, since the actual learning may be assessed more accurately through observation. However, when students are observed, the observer will be looking for particular behaviors (products) that indicate that the desired learning has taken place. If a student is evaluated while in the process of performing a specific selection of music, the emphasis in evaluation is generally the product which is heard by the evaluator. Aspects of process used by the student become products at the point in the learning process where they are evaluated.

Although summative evaluation is evaluation of product, this does not mean that products are evaluated only at the end of a learning episode. As students go through a process, many products are generated. These products can be evaluated at any time, and these evaluations can determine how well processes are working. As Rowntree (1987) noted, the teacher "cannot wait for the product to give him insights but must seek understanding while it is in the process of production -- intervening, where necessary, and with guidance"(p.138). How successfully students use various processes can be inferred at different points in the development of a product. In addition, the extent to which students engage in various processes and the efficacy of these processes can be observed through use of the various sources of evaluation data, particularly, observation.



In evaluation of process, it should be realized that students are active participants in a learning environment. They are doing things. In order to do things, students must have various skills. They must be able to interpret data, express themselves, participate in a group, do research, and have various psychomotor skills. Such skills are usually evaluated to ensure that students are acquiring those that will enable them to become active learners. Students are doing things which imply more of an open lesson approach to teaching as in case studies, role playing, document analysis, and problem solving. Students learn how to find out on their own, how to learn on their own, and how to make decisions for themselves.

Process is also designed to develop the affective objectives where students can begin to consider their own personal values, and teachers can help them develop the capabilities and processes they need in order to generate their own values. In the case of affective objectives, there would be no specific product if the teacher avoided forcing a particular value system upon students. In fact, the process of having students experience activities and having them develop their own values might be the product that the teacher is trying to achieve. However, the real emphasis is the process in which students examine issues and internalize their own value systems.

In evaluation of product, students' performance is assessed in relation to knowledge objectives. The recall of knowledge involves students remembering,

and recalling what has been presented. Knowledge may be broken down into three basic types: knowledge of specifics (terms and facts), knowledge of ways of dealing with specifics, and knowledge of universals or abstractions in a field. Knowledge of facts involve memory questions such as "Name two sources of water pollution." Knowledge of ways of dealing with specifics could require students to know how to end a business letter. Knowledge of universals could involve knowing what the concept "area" means in mathematics. Teaching and evaluating for product may also involve higher level cognitive skills. Since these higher-order skills require the presence of the lower levels of the taxonomy, the knowledge or product must be acquired if the higher-order skills are to be attained.

Objectives or goals have to be considered in light of their implications for evaluation. What is needed is a plan as outlined in **Appendix E**, that matches the objectives and the related specific learning outcomes to sources of evaluation data. Any evaluation effort, if it is to consider both product and process orientations, has to use a wide variety of data sources. In making a decision on appropriate methods of evaluation, the emphasis on the process-product continuum should be evident from the stated objectives.

The debate on the process-product orientations in the teaching-evaluation cycle may never be fully resolved. Although products have some intrinsic value, we should be equally concerned with the process that students go through

as they learn. Rather than simply learning to produce a product, students should learn in such a way that they develop a large measure of independence, and that they learn to think and make decisions for themselves rather than depend upon someone else. Griffiths (1987) advanced the notion that evaluation in science must "go beyond the level of simple recall and take into account the fact that the programme is process oriented" (p. 37). This principle clearly implies that even in subjects where there is a specific product, it is still very important to give due consideration to the process. Instructional approaches should be carefully selected to determine how the product and process outcomes defined by the objectives are to be achieved. It can be argued that the mere learning process implies intellectual growth or the acquisition of new knowledge, and hence a product orientation. In that sense, product is always present, so the debate can focus only on the degree of emphasis.

Differentiated Evaluation

It is sometimes felt that unless teachers evaluate all students in a classroom in the same manner and use the same sources for data, an element of unfairness enters into the evaluation. Such a view penalizes students with special needs who require modifications in the evaluation process if their educational growth is going to be assessed accurately. Each student is unique. Each student has strengths,

abilities, and areas of relative weakness. Recognition of these strengths and weaknesses, as well as knowledge of individual learning styles, should be the basis for modification of both instruction and evaluation.

In the process of evaluating *any* student, including the gifted, mentally handicapped, physically handicapped, emotionally disturbed, visually or hearing impaired, the evaluation method should not negatively affect a student's ability to demonstrate achievement of objectives. Just as we could not evaluate a hearing impaired child solely on the basis of an oral evaluation, we should not evaluate a child with a disability that impairs writing ability solely on the basis of written tests

In any evaluation process it is important that teachers evaluate only what they set out to evaluate. As Hutchison and Wonacott (1984) pointed out: "If a student fails a test because he or she cannot read it, you haven't tested his/her knowledge of the technical content - only his/her reading skill"(p.8).

The basic tenet of this section is that evaluation procedures should provide for individual differences among students with special needs. It follows, therefore, that evaluation procedures should be based on the stated objectives but modified according to the unique needs of the student. This is true for students in regular classes, in special education resource rooms, or in specialized programs of any nature. It is as true for special-

needs students following the standard program of studies of a particular grade as it is for those for whom an individualized program has been developed. It applies to school evaluation as well as public examinations, but the onus is on the school to request a modified evaluation. However, modifications for public examinations refer only to modification in one source of evaluation data (test) and apply to the cognitive domain objectives.

In the evaluation process, the teacher is trying to determine if students have achieved the objectives that have been established. If the objective is to determine whether students in a grade three social studies class understand the concept of community, methods for evaluation of this objective must be decided upon. The objective may be evaluated differently for some students. Some students may demonstrate understanding by answering written questions concerning the concept. Students with poor writing skills may use a map to explain the concept verbally. Some students may do a research project on communities, while others may demonstrate understanding of the concept through the use of a collage or poster. In all cases, students are showing that the objective has been met. What has changed is the method of evaluation.

If teachers want to find out what students know or can do, it is essential to ensure that the evaluation method itself not impede that process. It follows, then, that

modification of the evaluation process is desirable and often necessary. When evaluation procedures are modified, parents or guardians, as well as students when appropriate, should be involved in deciding the most appropriate modifications. Written reports should indicate the nature of evaluation modifications.

Differentiated evaluation has long been considered desirable by special education teachers, and writers in the field of special education. However, many students have special needs so modification of the evaluation process is legitimate and educationally acceptable for children of all ages and abilities. Hutchison and Wonacott (1984), suggested that **different methods be used whenever students' needs interfere with their ability to complete a particular evaluation.** The following is a compilation of evaluation modifications from many sources including the practices of good teachers.

- Skills or knowledge can be illustrated through demonstration. Students who have difficulty reading can demonstrate a knowledge of safety rules in an industrial arts class by correctly operating a power tool. Students who have difficulty writing well can demonstrate the understanding of condensation by drawing a diagram. Students who have weak verbal skills may be able to construct a collage or poster illus-

trating the concept of friendship in a family life course.

- Oral evaluation techniques can be used. Students who read well but have problems with written expression can use a tape recorder to answer questions that require more than a short answer. Students with a reading disability can have questions read to them and can dictate or write the answer. Students who are strong orally can give an oral report instead of a written one.
- Extra time can be provided for any evaluation activity. Students who have difficulty processing information may need more time to complete an examination. Students who are easily distracted may need to do tests or assignments in several smaller sections and in different sessions.
- Instructions on projects or tests may need to be clarified or stated in simpler terms. Students who have difficulty following directions may need monitoring during evaluation activities to ensure that they know what is expected. Often such students may be able to perform the required task but are confused by directions. Students with a weak, short-term memory may need directions stated one at a time rather than three or four together.
- The amount of work to be done can be reduced. Students who work more slowly than others may be able to complete only three division problems instead of five. Students who would have difficulty completing a full test may be able to complete only core questions. These questions would be indicated before the test is administered.
- Consideration should be given to not deducting marks for mistakes in spelling, grammar or handwriting unless these elements are being directly evaluated. Students with difficulties in the area of spelling may misspell some words on a science test but they may be able to approximate the correct spelling well enough to demonstrate knowledge of the science material. Students may be able to use a word processor with a spell-checker for stories, essays, or projects.
- Student performance may be video-taped and used to allow students to participate in the evaluation process. Students who have trouble understanding the concept of marks, grades, or comments in the evaluation process may benefit from seeing their own performance and from being encouraged to self-evaluate.
- Tests, examinations, and assignments can be modified to compensate for student weaknesses. Students who have problems with organization may need a formatted answer sheet. Students who have perceptual problems may need clearly-typed material to work with.

Handwritten tests may cause unnecessary difficulty for such students. Students whose reading ability is well below grade level may need questions modified to a lower reading level as illustrated in Figure 3.

large gymnasium. Alternative arrangements such as a smaller, more-isolated location can be considered. Students who are unfamiliar with the type of questions and the expected answer format may need some practice

Original Test Items	Modified Items
For each item circle the letter of the correct response. 1. Which is the correct procedure for wire-splicing? (a) connecting two pieces with a marrette (b) interlacing separate strands laid at right angles (c) joining separate strands with a screw-on mechanism (d) soldering two pieces with appropriate tools 2. What is the functional purpose of an awl? (a) to delineate the curvature of a boundary (b) to increase a circumferential boundary (c) to test the penetrability of a fabricating material (d) to transfer dimensions to a replacement fabric	Read each item below and circle the letter answer. 1. How is a wire-splice made? (a) connect with a screw (b) join with a marrette (c) twist the strands together (d) use solder and iron 2. What is an awl used for? (a) to determine the strength (b) to draw a line (c) to enlarge a hole (d) to mark correct measurements

Adapted from Hutchison, L., & Wonacott, M., (1984).

Figure 3: Modification of Test Items

- Anxiety can be a factor affecting a student's ability to attend to the task at hand in an evaluation process. Excessive student anxiety can be reduced. Easily distracted or nervous students may be adversely affected in an unfamiliar situations, such as writing tests in a

before the test is administered. They may need reassurance and monitoring during the evaluation process. Students who have had negative experiences with testing situations may be taught test-taking strategies and relaxation techniques.

- In evaluating students who are very bright or gifted, the practice of simply assigning more questions or extra work may not be the best idea. Instead of being required to provide one solution to several problems, they could be encouraged to provide several solutions to the same problem. They may be given assignments that require more sophisticated research skills, or be required to use the higher-order cognitive skills. Contracts can be used to specify expectations for units of work and the evaluation scheme. Self-evaluation of independent activities can be encouraged. Credit can be awarded for special skills and expertise. Independent projects can be used to demonstrate knowledge and skills acquired outside the prescribed curriculum.

Evaluation Issues in Effective Teaching Research

Early research tried to establish links between achievement and such personal characteristics of teachers as leadership, intelligence, appearance, and enthusiasm. This type of research produced some consensus on the desirable characteristics of teachers, but established no links between those characteristics and achievement. Consequently, the direction

of research efforts shifted toward the identification of those teacher behaviours (processes) which could be linked to gains in student achievement (products or outcomes), as measured by norm- or criterion-referenced tests. This type of research is generally referred to as process-product or process-outcome. Many of the findings of process-product research relate directly or indirectly to the practices used by teachers in evaluating student achievement.

Since process-product studies examined teacher effects upon achievement in the basic cognitive skill areas of reading and mathematics, their findings were not generalizable to development in other domains. Thus, the findings of these studies should never be regarded as a recipe for education, because education concerns itself with the development of the total person, not just cognitive ability in the basic skill areas.

In the remainder of this chapter the findings of selected studies dealing with monitoring and feedback which have implications for student evaluation are presented.

Monitoring

A model of instruction that predominates in many classes where basic skills are being taught is that of presentation-monitoring-feedback. Following a presentation, students often do seatwork. It is while students are engaging in

independent work that teachers engage in a particular type of formative evaluation - that of monitoring. Monitoring is a critical function of classroom teachers. Anderson, Evertson and Brophy (1982) pointed out that it is not enough for teachers to provide appropriate assignments and ensure that students know *what to do*. They must ensure that students know *how to do it*.

Many studies provide support for closely monitoring students. Some of the pertinent findings from specific studies include:

- Students who work independently under teacher supervision made greater academic gains than those who were expected to learn largely on their own. (Stallings & Kaskowitz, 1974)
- The most effective teachers were those who monitored seatwork carefully. (Tikunoff, Berliner & Rist, 1975)
- Highest academic achievement was found in classes where teachers monitored seatwork. (McDonald & Elias, 1976)
- The most successful teachers moved around the classroom during seatwork time. (Brophy & Evertson, 1976)
- The more successful teachers were those who monitored their students and had frequent, short contacts with them during seatwork times. (Emmer, Evertson & Brophy, 1979)

An issue which is closely related to the monitoring function is that of *accountability*. Brophy and Evertson (1976) discovered that in high-achieving classes, students were held accountable for careful and complete work. Their work was always checked, and additional instruction or assignments were provided where necessary. In classes of low socioeconomic status (SES), it was especially important for the teacher to see that all assignments were completed. In a study of fourth-grade classrooms, where teachers in the treatment group alerted students to the fact that their work would be checked at the end of the period and then promoted accountability by checking it, treatment groups outperformed control groups on both criterion-referenced and standardized achievement tests. (Good & Grouws, 1975)

Very often, teachers monitor student progress during lessons through questioning. Research indicates that there is no best way to select students for questioning. Findings vary according to grade, SES, and size of group. In the early grades, it is important that all students participate in answering questions. This is especially true of small groups. In such groups, it is advisable that students take turns because this ensures that all students participate about equally (Brophy & Evertson, 1976). Assertive students should be prevented from dominating, and reticent students should be required to participate.

Although it appears that by Grade 7, students can learn effectively without overt participation in lessons (Hughes, 1973), teachers of higher grades should be

careful not to interact with the same students all the time. Brophy and Good (1986) suggested that volunteering answers should be encouraged, and pauses should follow questions to give students time to think. Non-volunteers should be called on frequently, especially when it is likely that they will respond correctly.

Teachers often have to contend with students calling out answers. Where call-outs are prevalent, it can be very difficult to monitor individual students. Research showed that the effect of student call-outs varied with SES level. They were correlated positively with achievement in low SES groups but negatively in high SES groups. In a synthesis of the research findings, Brophy and Good (1986) suggested the following:

When most students are eager to respond, teachers will have to suppress their call-outs and train them to respect one another's response opportunities; however, when most students are reticent, teachers will have to encourage them to participate (which may include accepting relevant call-outs). (p.363)

In studies of primary, elementary, and junior high students, Brophy and Evertson (1976) and Emmer et al. (1979) concluded that it is important for teachers to work for any response at all from uncommunicative students. The response of students who give incorrect or incomplete answers should be improved by the teacher.

Feedback

Many researchers have investigated the effects of providing feedback to student responses. Denham and Liberman (1980), found that the frequent provision of academic feedback was a positive correlate of achievement. Stallings and Kaskowitz (1974) discovered that almost anything connected with the pattern of questioning-response-feedback correlated positively with achievement. Stallings, Cory, Fairweather, and Needels (1977), noted that in Grade 3 classes which made the greater academic gains, there was more instruction, more questions, and more feedback. In a study involving secondary students, one of the behaviours of teachers in the treatment group was that of reacting to student responses (Stallings, 1980). Although differences in levels of achievement between the control and experimental groups were not statistically significant, students in the experimental groups gained an average of 6 months more in reading achievement. In their study of small groups of Grade 1 students, Anderson et al. (1982) concluded that students should receive clear feedback about the correctness of their performance. They suggested that, unless it is very obvious, correctness should always be acknowledged.

Quality of feedback may also affect student achievement. Hughes (1973) concluded that teacher reactions to student responses can influence motivational levels. Enthusiastic responses are more motivating than matter-of-fact reactions.

In a correlational study involving Grade 4 students, Good and Grouws (1977) found that the teacher behaviour of providing process feedback (explaining the steps involved at arriving at an answer rather than simply providing the correct answer when a student failed to respond or responded incorrectly) correlated positively with achievement. Stallings, Cory, Fairweather and Needels (1978), in a study of secondary students, found that the provision of support and corrective feedback when students did not respond correctly was a positive correlate of achievement. In a study of Grade 6 students, Clark, Gage, Marx, Peterson, Stayrook and Winnie (1979) concluded that it is very important to provide feedback to student responses and to attempt to improve responses that are incomplete or incorrect.

An element that may be part of the feedback stage is *praise*. Researchers have studied its effects rather extensively. In a study involving Grade 6 students, Dunkin (1978) found that teacher praise of student responses was positively correlated with achievement. Emmer et al. (1979) discovered that teacher praise was relatively frequent in high achieving junior high classes. Praise was also a positive correlate of achievement in the study of secondary students by Stallings et al. (1978). McDonald and Elias (1976) concluded that positive feedback and praise were positively correlated with achievement in second grade reading and fifth grade mathematics. Stallings and

Kaskowitz (1974) noted that praise for correct responses and good work tended to correlate positively with achievement, but more in first than third grade, more for math than for reading, and more for low than for high-ability students.

Brophy (1981) reviewed much of the literature on teacher praise. He concluded that quality was more important than frequency. The following guidelines were synthesized from the research:

- Praise should always be contingent upon good performance, rather than being delivered randomly or unsystematically.
- It should specify the particulars of the accomplishment rather than be global.
- It should show spontaneity, variety, and other signs of credibility with clear attention to the student's accomplishment rather than being delivered with a tone that suggests a conditioned response made with minimal attention.
- Actual work and effort, not mere participation, should be praised.
- It should provide information to students on their competence or the value of their accomplishments rather than giving no information to students about their status.
- It orients students toward better

appreciation of their own task-related behaviour and thinking about problem-solving rather than orienting them towards comparing themselves with others and thinking about competing.

- It uses student's own prior accomplishments rather than the accomplishments of peers as the context for describing present accomplishments.
- It is given in recognition of noteworthy effort or success at tasks which are difficult for the student rather than being given without regard to the effort expended or the meaning of the accomplishment.
- Praise attributes success to effort and ability rather than to ability alone or to external factors such as luck or low task difficulty. This implies to the student that similar successes can be expected in the future.
- Students should be praised for effort, not for pleasing the teacher or winning an award.
- Effective praise focuses students' attention on their own behaviour not on the teacher as an authority figure.
- Praise is delivered after the process is completed, not while it is ongoing. (Obviously, a distinction is being made between praise and encouragement here.)

It should be pointed out that praise, by itself has little value. It is valuable only in the context of effective teaching and good classroom management. Coker, Medley and Soar (1980) found in a study of Grade 1 and Grade 2 students that achievement was negatively associated with poor classroom management and overemphasis on praise and rewards. Thus, praise for praise sake or the use of praise as a teaching strategy may be counter-productive.

The question of how teachers should react to inadequate responses arises, especially where incorrectness or incompleteness is due to lack of effort. Tikunoff et al. (1975) noted that the most effective teachers were polite and pleasant in their interactions with students. The atmosphere in their classrooms was characterized by mutual respect and these teachers were less likely than the less effective teachers to ignore, belittle, harass, shame, put down, or exclude their students. Brophy and Evertson (1976) pointed out that criticism for poor response or poor work correlated positively with achievement in high-SES classes only.

Anderson et al. (1982) used the results of process-product studies to develop guidelines for small group instruction in beginning reading. They have suggested that correction, not criticism, be used when students respond incorrectly. When it is necessary to criticize, teachers should be very specific about what is being

criticized and about desired alternative behaviours.

Generalizations From Research Findings

The generalizations which teachers can use as a guide in the evaluation process, include:

- Once seatwork is set, teachers should monitor students carefully, preferably by moving around the classroom.
- Seatwork should always be checked.
- Teachers should ensure that all assigned work is completed.
- When teachers monitor classes through questioning, all students should be given the opportunity to participate. Although this is less critical at the intermediate and senior high, it is still important. Teachers should pay particular attention to having reticent students participate.
- Unless the correctness of an answer is very obvious, teachers should always provide feedback to students on the adequacy and correctness of their responses. Feedback should be process-oriented and corrective when responses are inadequate or incorrect.
- Feedback should be provided as quickly as possible. Students should receive grades on examinations within 2 or 3 days after they are written.

- When students are praised for specific accomplishments, they should receive it in a manner which recognizes and shows the value of their accomplishments and encourages them to compete with themselves rather than with others.
- Even when students give less than a good effort, their dignity must be respected. No student should be ignored, belittled, harassed, shamed, or otherwise treated with disrespect.

Summary

This chapter has dealt with the various aspects of the evaluation process. The first section distinguishes between three types of evaluation. Pre-instructional evaluation involves the gathering of information on what students bring to the learning environment. Formative evaluation refers to ongoing evaluation during the instruction and learning process, while summative evaluation takes place at the end of a learning episode to grade, certify or select students.

Regardless of the type of evaluation that is used, the initial step in the process is the setting of goals or objectives. It is not enough that teachers "know" what they plan to teach; they should clearly articulate objectives and communicate them to students. Objectives should guide the teaching-learning process, and become the basic reference for student evaluation.

They should include each domain of learning and represent lower-order and higher-order levels within each domain.

This focus on the process of evaluation has explored the process versus product approach to instruction and evaluation, the significance of differentiated evaluation, and the various findings in the effective teaching research as they pertain to student evaluation. Process and product approaches to instruction are interrelated elements, both in teaching and in evaluation. A balanced orientation gives significant treatment to both the process of learning and the actual content which students acquire. Differentiated evaluation highlights the child-focused emphasis of an effective approach to evaluation. Consideration should be given to a student's level of physical, sensory and mental functioning, since the teacher should consider not only *what* is being evaluated, but *who* is being evaluated. As for the effective teaching research, it provides a sound rationale for a number of significant evaluation issues that should become an integral part of the teaching-learning process.

CHAPTER III

SOURCES OF EVALUATION DATA

Data may be obtained from homework, learning journals, observation, personal interviews, projects, self-evaluation, tests and work sample files. All of these sources, listed in alphabetical order, are discussed in this section of the document.

Individual teachers may have difficulty in deciding which source is the most appropriate in a particular situation. No source is likely to provide all the required information because all sources have limitations. Consideration must be given to the objectivity and accuracy of formal and informal observation, possible bias in self-evaluation reports, rating scales and anecdotal records, and the validity and reliability of tests. The source should be chosen in terms of the purpose to be served, rather than on the basis of its accuracy, ease of administration, or its objectivity. The essential question for a teacher to answer is "How can I best determine what I want to know about the student?". All other considerations are secondary. Thus, a combination of sources should be used.

Homework

When teachers are evaluating students, the results of homework assignments may provide worthwhile information for evaluation. In 84 experiments conducted between 1904 and 1984, a majority indicate that homework can increase

student performance on criterion referenced tests (Foyle & Bailey, 1986). Cooper (1989), after reviewing 120 studies of homework, concluded that homework can have a positive effect on the achievement of elementary and high school students.

Homework is a pervasive aspect of the student's life. Thus, its assignment and evaluation ought to be guided by basic tenets which are clearly understood by administrators, teachers, parents, and especially the students. The following description is aimed at assisting administrators and teachers in developing and implementing homework policies that will benefit the student.

Homework may be classified in terms of five types: (Foyle & Bailey, 1986)

Completion homework is assigned to students who do not finish work assigned in class. Such homework should not be assigned to students who do not demonstrate some degree of understanding.

Creative homework requires students to combine numerous skills and concepts to produce a required response.

Extension homework determines whether students can transfer acquired skills or concepts to a new situations.

Practice homework helps students reinforce skills. It should be limited to

material presented in class. It should not be assigned to students who do not demonstrate some degree of understanding. Otherwise, someone other than the teacher would have to provide instruction.

Preparation homework prepares students for subsequent lessons such as when teachers request students to read the next four pages in their social studies text. If this is the only guidance given, the maximum benefit will not be gained from the homework assignment. The students deserve preparation for such homework. A study guide could be prepared which requires students to give brief answers to literal comprehension questions. This provides a purpose for the reading. Care should be taken to deal with any difficult or new vocabulary where context is not strong enough for students to deduce meaning. This kind of guidance is especially important for the younger or struggling student.

Homework, like other sources of evaluation data, is only a means to an end. Its value is determined by how well it furthers educational objectives. If it does not, then it serves no purpose. For example, when homework is assigned to provide remedial instruction, giving the same blanket assignment to all students is hardly justified. Rather, the assignment of homework ought to provide for some individualization.

The amount of homework that should be assigned is a complex question. Some

obvious considerations are the students' age, ability, habits, and the instructional objectives. When students have several teachers, it is essential that these teachers develop a homework schedule. Otherwise, students may be assigned five or six hours of homework in an evening.

Early in the school year, homework expectations should be clearly described. For example, will work not completed in class automatically become homework? If a specific assignment is not made for the next class, will students be expected to read ahead in their texts, review previous work, begin or work on some project? Homework procedures ought to be as familiar as any other part of the classroom routine.

Homework should never be assigned as punishment. This too common practice results in students believing that homework is an unpleasant activity. As well, increasing the amount of homework to compensate for lost class time produces nominal, or even negative results.

There are numerous ways in which teachers may prevent homework problems:

- Develop a systematic, unvarying assignment schedule.
- Develop a standard method for giving assignments.
- Provide a framework for assignments so that students clearly understand the task.
- Be reasonable in estimating the amount of time required to complete a

task, and consider other obligations that students may have. Be flexible if students request extra time.

- Minimize student choices and decisions once they begin independent work.
- Ensure that assignments can be completed with available resources.
- Counsel students concerning work habits and methodology.
- Use other school personnel and resources to assist students.
- Assign and check work in small amounts.
- Encourage students to ask questions, and provide direction concerning questions they ought to be asking.
- Provide supervised time, if possible, for students to get started.
- Write comments on assignments. Praise and helpful criticism can produce positive results.

Teachers may find the following ideas helpful when giving homework assignments:

- Do not compete with student noise when giving oral instructions. Stop talking and wait for their attention.
- When possible, provide written instructions. Be sure to write clearly and keep a master copy. Should comprehension problems arise in completing assignments, place refinements on the master copy.
- When possible, choose assignments that are relevant to students' interests. Given a broad topic, students may decide on a particular aspect that has legitimate educational purpose and will provide fun in addition to knowledge.

- When assigning major projects to be completed outside of school, display sample projects to clarify expectations and to minimize questions.
- Reserve specific times for regular assignments. For example, on Mondays students read the next chapter in their text. On Thursdays they hand in answers to review questions. Students will be less likely to forget assignments if a routine is established.

It appears that, whatever the homework purpose, assignments must be:

- regularly assigned
- clearly stated
- regularly collected.
- promptly and consistently graded.
- promptly returned.

Finally, the question of the role of the parent in homework arises. Often, parents can contribute greatly to their child's education through their involvement in homework activities. It is not, however, the responsibility of parents to provide the instruction that students need to do homework. Parents should provide support and encouragement. They should ensure that their children have a quiet place to do homework, that children's schedules contain homework time, and that children actually do assigned homework.

The effective teaching research indicates a positive correlation between achievement and the amount of time students spend attending to tasks where they experience a high success rate. When

students can do the assigned homework, there is an increase in the amount of academic learning time. Therefore, homework which is well-planned and for which students are well-prepared can have much the same effect on achievement as a lengthened school day.

Learning Journals

Students may keep journals or diaries in which they make daily entries about what they have learned. Sometimes more structure may be needed and teachers can specify subjects to be discussed in the journal. The teacher, with the student's consent, should read the comments to determine whether or not major objectives are being met. Since journals may reveal how students are processing what is happening in the classroom, journal entries may sometimes highlight misconceptions which should be corrected before the unit continues. Also, as Muir and Wells (1983) indicated, it can be quite revealing and useful to have students write about those things they have not understood in addition to what has been learned.

Observation

Observation helps the teacher get to know the student as an individual, because it takes into account those aspects of student performance or behavior that cannot be

easily assessed through information from other evaluation sources. Since observation is concerned with process as well as product, it is used to monitor what a student is attempting to do and what the student can do. For example, the process a student uses to cope with reading a difficult text will provide evidence of how that student is developing reading strategies. Observation should be widely used for evaluation in the primary school and should be an essential component of evaluation in all subsequent grades. *The Primary Mathematics Curriculum Guide* (Department of Education, 1988a) states that observation is the most important source of evaluation data in primary mathematics. *The Special Education Policy Manual* (Department of Education, 1988b) also emphasizes the importance of observation in identifying students with special needs. Language 2101, Advanced Writing 3104, and French 3200 are examples of senior high courses in which observation could be used as a primary source of evaluation data.

Formal observations must be carried out regularly to provide information on students' growth and development. The teacher must make informal observations as well as systematic, formal observations of individual student's physical, social, emotional, spiritual, moral, and intellectual growth. This enables the teacher to gain a clear picture of a typical pattern of behavior over a period of time in a variety of situations. Observing a student in a variety of situations enables the teacher to note patterns in the student's behavior and to make decisions

about the amount of assistance needed to revise instruction or to adjust activities.

Whenever a student reveals special needs that may interfere with learning, a more precise identification of these needs should be made with the help of specially trained personnel. (See *Special Education Policy Manual* for a more indepth discussion)

Observation Guidelines

A systematic approach to observation is required in order to maintain objectivity and to make the process effective and reliable. It is most effective as a method of evaluation when the following guidelines, developed with reference to *Look Hear* (Board of Education for the City of



North York, 1983), are taken into consideration:

What to observe: Consider the development of the whole student.

- Physical
- Intellectual
- Social
- Emotional
- Moral
- Spiritual
- Esthetic

When to observe: Plan to observe on a regular, systematic and frequent basis.

- Observe each student in a variety of situations.
- Plan to observe daily, weekly, and long-term.
- Observe the student in different situations.
- Record development without interacting with the student.
- Record development while interacting with the student.

How to observe: Be as positive as possible. While negative aspects cannot be ignored, look for successes rather than failures.

- Develop a realistic plan that concentrates on a small number of students.
- Focus on one or two specific criteria at a time.
- Wherever possible, use audio and video tapes.
- Use the student's own language when making notes.

- Develop a checklist to aid in recording data.

Use of observation data: Allow time for the systematic review of, and reflection upon, recorded observations:

- Determine which methods and materials each student will need to develop the performance expected.
- Confer with colleagues.
- Confer with students.
- Report to parents.

This does not mean that the teacher cannot use informal observation. In formal situations, tasks are prescribed and the teacher observes and evaluates. Emphasis is on what students **can** do. During informal observation, the teacher does not direct the student but remains unobtrusive. This type of observation provides information on what students do spontaneously. Informal observation may yield much useful information, especially about affective behaviors. Results of both types of observations should be recorded.

Observation Recording Devices

Gronlund (1985) states that impressions gained through observations are apt to provide an incomplete and biased picture unless an accurate record of observed behavior is kept. Therefore, when teachers evaluate students through observation, a record must be made either while the observation is being conducted or as soon as possible thereafter. Records should be manageable. Record only what

is useful. However, the use made of records is as important as their maintenance. Records can be used to aid:

- Teachers in evaluating individual students.
- Students in making self evaluations.
- teachers in improving programs and instruction.
- Teachers in communicating with colleagues and parents.

Checklists. A checklist, as seen from the example in **Appendix F-1**, is a list of specific observable behaviors that can be prepared by the teacher or adapted from commercial sources. The checklist permits the teacher to note the presence or absence of desired attributes. Teachers should make as much use of checklists as possible, because they provide immediate descriptive information. They allow for on-the-spot evaluation and, in many instances, rapid feedback about the adequacy of the performance or product. Checklists should be regarded as guides to developmental milestones and not as a sequenced list of skills. One of the most important uses of checklists is as a basis for anecdotal records.

The following are suggestions for constructing checklists developed with reference to *Making the Grade* (Board of Education for the City of Etobicoke, 1987) and *The Primary Curriculum Handbook* (Department of Education, 1987).

- Items should be clear and precise.

- Items should reflect the instructional objectives for the unit of work.
- Keep checklists short to allow the observer to focus on specifics.
- Design checklists so that they can be used as a basis for discussion with students.

Anecdotal records. Anecdotal records allow for the recording of more indepth observations and information. They are also excellent sources of information for conferences with parents and students (See **Appendix F-2**). The anecdotal record must be based upon what actually was observed, since it is a description of some specific episode. A collection of such descriptions recorded over the school year provides the teacher with a documented account of a student's progress.

During observation, it is not always convenient to make detailed anecdotal records. Therefore, when the observation reveals important data about the student's development, the teacher can make notes of circumstances, dates and names. After the observation has been conducted, these notes can be used to write the actual anecdotal record.

The following guidelines for writing anecdotal records have been developed with reference to Baskwill and Whitman (1988) and the *Primary Curriculum Handbook* (Department of Education, 1987).

- Always note names and dates for later reference.
- As soon as possible after the observations, use notes to record more formal reports.
- Always maintain objectivity when compiling anecdotal records.
- Always note strengths as well as weaknesses.
- Have a system of maintaining these records for the student's permanent file.

Rating scales. Rating scales, as outlined in **Appendix F-3**, are similar to checklists, but they also include a measure of performance quality. As an evaluation instrument, a rating scale is useful in determining a student's strength and weakness, and in planning a suitable program. Ratings should be criterion-referenced to ensure that a student's performance is stated in relationship to specific course objectives and not in relation to other students in the class.

Projects

Projects are formal assignments related to the curriculum. They may be assigned to a student or group of students, and they usually involve some type of research or development such as constructing models, preparing reports, or producing plays. While essays may share many of the

characteristics associated with the project, they are considered to be a separate entity. In this section, the discussion has been limited to formal projects, and is not meant to apply to the informal projects that even primary students sometimes do.

It is **not** generally appropriate to assign formal projects before Grade 4. In Grades 4, 5, and 6, projects should be completed at school with considerable teacher assistance. Beyond Grade 6, students should exercise increasing degrees of independence on assigned projects.

Some advantages of projects are that they:

- Provide students with an opportunity to formulate self-motivated questions and answers.
- Allow students to demonstrate individual or group initiative.
- Provide students with an opportunity to use concrete materials in illustrating their ideas/talents.
- Present an opportunity for assessing skills not evident with written or oral tests.
- Can be assigned depending on ability levels within a given class.
- Allow for student choice in pursuing individual interests in demonstrating otherwise hidden talents.
- May be used as a teaching tool, if completed projects are used for

classroom instruction or presented by the students.

- Can provide the opportunity for idea exchanges.
- May provide an alternative for students who have difficulty with writing.

Projects also have some disadvantages.

- Projects may be enhanced or diminished by the home environment.

If a parent completes most of the project, then the student may have learned little. Yet, such a project may be given a higher grade than one which was totally completed by a student. However, if a family member has offered advice or assisted to a limited degree, then the learning experience might have been quite valuable or even enhanced by the contribution of the family member.

- There are important questions in terms of project evaluation that cannot be easily answered.



- Will the teacher be able to determine how much of a project has been done solely by the student?
- Is it possible to establish and apply objective marking criteria?
- Can the same evaluative criteria be used to assess the full range of projects?
- A considerable amount of time will be required by the teacher for organizing, marking and planning. Where projects contribute to a final grade, a considerable amount of time may be required by students for a small portion of the final grade.
- Competition may create peer comparisons which are harmful to the self-esteem of some students.

Guidelines for Assigning and Grading Projects

- Be sure that students understand what the project is and that they have some basic methodology formulated. That is, students must be taught how to complete a project *before* it is assigned.
- Give students ample opportunity to discuss a general topic. From this, they can select a smaller, more manageable portion as their own project. For example, the topic "School Days in the 1930's" is too broad, but "School on Woods Island in the 1930's" is quite manageable for most students.
- Be sure that projects are not too difficult so that students may derive personal satisfaction from doing them.
- Have students start projects in the classroom, especially in Grades 4 to 10.
- Be sure all students are aware of the teacher's expectations regarding the project *before* it is assigned.
- Provide students with a written outline of the project. This should include the following:
 - the objectives of the project.
 - the specific terms of the project.
 - the due date (or deadlines for phases of the project).
 - evaluation criteria (including a marking scheme).
 - suggestions concerning sources.
- If possible, resource centre personnel or other people who might be involved should be contacted about availability of project materials, unless the teacher prefers that students locate information themselves. This will help develop skills in using references.
- Continuous consultation should be maintained with students about their plans and how they are progressing. This ensures that projects do not

become unmanageable or unrealistic in scope or time. It also ensures that students do not get off track.

- Avoid assigning the same project topics in one or more succeeding grades or levels.
- Provide marks or comments regarding the student's completed projects on a separate sheet of paper. *Do not* place any markings on the project itself. This allows the work to be displayed. However, work should *not* be displayed without consent of the student, and never for comparison or competition, unless it is assigned for that purpose.
- Allow students to generate their own topics for projects, and consider their suggestions for development.
- If possible, encourage students to choose a project rather than an essay.
- Allow students to volunteer for group projects with the understanding that each member receives the same grade for the completed work.

Evaluation Criteria for Projects

Two aspects of projects to be evaluated are process and product. Process refers to the activities which lead to the completed project, or product.

Process evaluation. Information gathered by considering the following questions can assist the teacher in preparing anecdotal comments for diagnostic and interviewing purposes:

- Was the assigned task clearly understood by the student?
- Did the student indicate a positive attitude toward the assigned task?
- Did the student indicate a willingness to be involved in the assigned task?
- Did the student undertake extra work voluntarily?
- How much personal responsibility was assumed by the student?
- Was the student open to suggestions from teacher, peers, and other sources?
- During teacher-student consultations, peer discussion groups, or progress reports:
 - Was the student always aware of the purpose of the project?
 - If working in a group, did the student encourage others toward completion of the project?
 - Did the student remain topic-oriented throughout the project?
 - If working in a group, was the student sensitive and attentive to needs of other members?

- Did the student demonstrate the thinking, research, communicative and group participation skills necessary for locating and using information?

Product evaluation. The following questions can help the teacher develop criteria for evaluating the project:

- Did the completed product clearly indicate that the student understood the task?
- How well were all the required elements incorporated?
 - Were they complete?
 - Were they relevant?
 - Were they accurate?
- Was the project organized?
 - Was there unity of elements?
 - Was it organized in logical and sequential pattern?
 - Did the project have originality in terms of ideas and presentation?
 - Did the project have a positive impact, visual or conceptual?
 - Was there an appropriate emphasis?
 - Did the student use support materials?

An example of a project assignment sheet is contained in **Appendix G**. It provides a framework for assigning a project. In addition, discussion time should be allo-

cated, both prior to and during project completion.

The literature suggests that projects are primarily a learning activity, *not* an evaluation activity. In practice, most educators know that the effort applied by students is usually proportional to the value assigned to a project. Many courses such as Industrial Arts and Home Economics are designed specifically to accommodate projects. Other courses such as Literature, Culture, and the sciences, lend themselves, very easily, to a project approach. Thus, in the years prior to junior high, it may be unwise and unnecessary to assign formal projects with evaluation as a motivational factor. However, in subsequent years, the project may become an essential factor in student evaluation.

When a grade is assigned to a project, it should be derived from a consideration of process and product factors mentioned previously.

Self-Evaluation Reports

Self-evaluation provides an opportunity for students to examine their own progress and to analyze their own strengths and weaknesses.

The self-evaluation report, which could be based on the example in **Appendix H**, is a very effective technique for teaching

students acceptance of the responsibility for their own learning. Self-evaluation helps students to think about what they are doing and what they need to do. When students are aware of their own weaknesses, motivation to improve will likely increase. As Andrews et al. (1986) indicated, to evaluate their own work effectively, students need to be trained to report:

- what was done;
- how it was organized;
- use of resources, and
- what improvements could be made.

Teacher-Student Conferences

The teacher-student conference, like direct observation, is a very practical way for a teacher to develop a full picture of a student. It provides an opportunity for the teacher to focus on an individual student.

Some students may behave much the same in an interview as in other activities, but a different side of a student may be revealed in an interview. The shy student who usually works alone may be quite verbal when given one-to-one attention by an adult. The student who is boastful and assertive with peers may be insecure and uncomfortable with an adult. Interviews must be carefully focused and preplanned. The student should be at ease and encouraged to be open and honest. In the

interview situation, judgmental remarks on the part of the teacher should be avoided.



When planning and conducting interviews, teachers may find the following guidelines, mostly from the work of Phinney (1982), very helpful.

- Prepare clear, concise questions.
- Create a relaxed environment.
- Concentrate on the student's information.
- Allow the student sufficient time to pause and think.
- Allow students to do most of the talking.
- Record the results of the interview.

The Special Services Division of the

Department of Education has produced a videotape that should be very helpful to those who wish to use the teacher-student conferences as a source of evaluation data. *Math Match*, (Department of Education, 1989) also contains a comprehensive discussion of parent-teacher conferences.

Tests

The source of evaluation data most widely used in the classroom, especially at the intermediate and senior high levels, is the teacher-made test. In fact, it is generally felt that testing is abused, because it is often the only source of data for important educational decisions related to performance of students. Thus, teachers need to have a clear understanding of types of test, test construction, test administration, validity and reliability. They should realize, however, that a test is only one source of evaluation data. For a reasonably valid and accurate evaluation, other sources of data should be included in the process.

Types of Tests

Norm-referenced tests. These tests are constructed to determine where an individual or group stands in relation to a comparison or norm group. They are not designed for the specific purpose of determining what a student can or cannot do. However, a careful analysis of the performance of a student or class can

provide information that can be extremely valuable in the teaching-learning process. Norm-referenced tests, are usually developed commercially and some degree of mismatch between what is taught and what is tested is to be expected.

Gronlund (1985) emphasizes additional characteristics of norm-referenced tests. He states that they cover a large domain of learning tasks, with just a few test items for each task. They favour items of average difficulty and emphasize discrimination among individuals.

Criterion-referenced tests. These tests are designed to determine where an individual stands in relation to a set of objectives. The rating for a student is in relation to the stated objectives, and is not affected by the performance of the group. Criterion-referenced tests focus on a limited domain of learning tasks, place emphasis on determining what individuals can or cannot do, and match item difficulty to learning tasks. (Gronlund, 1985). Teacher-made tests are criterion-referenced, and are administered to measure the performance of students in relation to specific objectives.

Validity

Although there are four different types of validity, classroom teachers are most concerned with content validity. Simply put, content validity is concerned with the degree of match between test items and course objectives. Empirical procedures, requiring the ratings of judges may be

used, but they are not necessary. Teachers may assess content validity by examining each item of a test and comparing it to the objective it is meant to measure. If it is judged to measure the objective it is meant to measure, it may be regarded as valid. If not, it may be regarded as being invalid.

Validity cannot be assessed in absolute terms. An important consideration in any assessment of validity is the use to which test results are put. As Cronbach (1971) has noted:

A single instrument is used in many different ways - Smith's reading test may be used to screen applicants for professional training, to plan remedial instruction in reading, to measure the effectiveness of an instructional program, etc. Since each application is based on a different interpretation, the evidence that justifies one application may have little relevance to the next. Because every interpretation has its own degree of validity, one can never reach the simple conclusion that a particular test "is valid". (p.447)

Reliability

Reliability is concerned with the amount of random fluctuation there would be in individual test scores if a test were given over and over. The requirement that test results be stable over time is secondary to the requirement that it be valid. It is

noteworthy as well that a test can be reliable without being valid. People who want to tell if they are losing weight may use an inaccurate scale for that purpose. The scale will tell if a person is losing or gaining weight because it is reliable in providing measurements of weight. However, the reading on the scale is not a valid one.

There are many ways of increasing reliability, but these are usually restricted to the domain of norm-referenced testing. The only consideration for teachers with respect to reliability is that tests not be too short. There should be five or six selected-response items testing each objective, for example. If true-false items are used to test the same objective, there should be more. In a test composed of constructed-response items, reliability is enhanced when there are clear guidelines for the marking of answers.

Test Construction

In addition to validity and reliability, item type must be carefully considered when constructing a test. The major decision is the relative percentage to be assigned constructed-response items (essay and completion objectives) as opposed to selected-response items (multiple-choice, true/false, matching). Subsequent decisions regarding the nature of the items, and the effort, planning and careful analysis during the construction and review phases, will determine the effectiveness of the test.

Constructed-response test items. These items force students to develop their own answers, as in the case of essay-type questions or completion questions. Questions that pose specific problems and require students to recall knowledge, organize it, and produce defensible conclusions are known as *restricted-response* items. Those which allow students to determine the length and complexity of their answers are known as *extended-response* items (Kubiszyn & Borich, 1984). While these items are usually easier to construct as seen from the guidelines contained in **Appendix I**, the subjectivity involved in scoring, and the time required for marking impose limitations. The general guiding principle is that constructed-response items should be used for objectives that focus specifically on the generation of ideas. In particular, extended-response items are considered more appropriate for the cognitive processes of synthesis and evaluation.

As a guide for teachers on the writing of test items, a list of directing words corresponding to each level of Bloom's taxonomy is provided in **Appendix A**. The list is not meant to be exhaustive, nor are the words meant to be confined to one category, and some can be used in categories other than those in which they are listed.

Good constructed-response questions are not easy to develop, but they are much easier to develop than multiple-choice

items. When groups are small, and more time is available for marking than for test development, the use of constructed-response questions may be quite defensible. Also, if examination questions are not to be reused for security reasons, it might be wise to develop constructed-response questions because a good constructed-response test takes less time to construct than a good selected-response test. Finally, if it is an objective of instruction to develop ability in written expression, it might be more appropriate to use constructed-response questions.

Constructed-response questions have some serious limitations. One of their biggest weaknesses is their subjectivity. This subjectivity makes it difficult to produce reliability in grading, because of extraneous factors such as writing ability of students. Also, less content can be sampled. Another problem involving constructed-response items is that students may be unsure of exactly what they are to do. Vague directing words such as "discuss" give the student very little guidance. The answers given to such questions are generally so wide-ranging that teachers may have problems in grading. Marking guides could never cover all possible responses to such vague questions.

Despite these weaknesses, constructed-response items are very common in educational testing. This is due to tradition and to the widely held belief that the use of constructed-response items has

a more beneficial effect than any other type of item upon student learning and student study habits.

Selected-response test items. These items require students to select an answer from a set of alternatives. The major advantage is scoring ease and accuracy.

In general, selected-response measures should be used for assessing most skills and objectives dealing with recall of knowledge. However, if properly developed, they may be used to measure higher-order cognitive processes. The most common types of selected-response items are true-false, matching and multiple-choice. All of these item types have various strengths and weaknesses and each has a set of guidelines as outlined in **Appendix J** that should be followed closely when items are being constructed.

Selected-response tests require much less time for students to complete and teachers to mark, and measurement becomes more efficient. Since more items can be included on a selected-response test, and scoring is less likely to be biased, reliability is likely to be higher than on tests requiring constructed responses. In practice, teachers may often find that, the best tests in many subjects will contain both selected- and constructed-response items. Selected-response items measuring higher-order cognitive processes are difficult to write. This is not to suggest, however, that all selected-response items should deal only with knowledge-recall objectives.

Good selected-response items do not permit correct responses on the basis of recognition alone. They require examinees to develop a basis for choice through creative, original thought (Ebel, 1972). The argument has been made that constructing an answer is no more complex or difficult than choosing one. In fact, studies have shown that scores on selected-response tests and scores by the same examinees on constructed-response tests measuring achievement of the same objectives correlate almost perfectly.

One point that should be considered is that selected-response items allow more time for reflective thought on the part of examinees. Reading requires less than 10% of the time that writing requires, giving students more time to think. The counter-argument to this point is that reading ability is a factor that always enters into test results. However, this would be no more of a confounding factor than writing ability on a constructed-response test.

It is often assumed that selected-response test items can measure achievement only on lower-order recall objectives. Although it is true that synthesis and evaluation cannot be measured adequately using selected-response items, higher-level cognitive processes, such as understanding, can be measured through the use of properly prepared items. For more detailed information and examples of lower and higher level selected-response test items, see **Appendix K**.

Test Construction and Administration

Early research claimed that it made little difference how items were organized on a test. However, in recent years, a number of researchers have found this to be incorrect. The following guidelines should help remove many of the problems associated with construction and organization.

- Items should be directly related to the objectives.
- Items should be clear and free from ambiguities.
- Items should contain vocabulary about two years below the grade level of students.
- Items should be grammatically correct and free from spelling and typing errors.
- Items should test information that students are expected to know and use.
- An item should *not* be based on the response to another.
- An item should *not* provide a clue for another item.
- Items should cite authorities for statements that might be considered debatable or based on opinion.
- Items should be written in a straightforward, simple manner.
- Items should use the simplest method for requiring a correct response.
- Test directions should be clear and complete.
- Items should be grouped by type.
- Items should be grouped according to instructional content.

- Items, within groups, should be arranged in order of increasing difficulty.

When administering tests, teachers should give students every opportunity to do well. Thus, a careful consideration of factors that may affect test results is important.

- **Physical setting:** Every effort should be made to provide an atmosphere conducive to the student's comfort. This should include ample space between students; proper lighting and ventilation; comfortable temperature and air circulation, and freedom from distractions occurring outside the testing room. Here, timing may help to avoid problems.
- **Psychological conditions:** Teachers can help reduce anxiety by adequately preparing their students. This is best accomplished by reviewing the content to be covered by the test and providing students with sample items. Tests should never be given as punishment.
- **Time:** Students should be given time to write classroom tests.
- **Directions:** All students should clearly understand what they are to do.

Test Practices

The discussion which follows provides the basis for the development of guidelines for testing.

Comprehensive examinations. When grades are assigned as numbers or percentages, a fair amount of testing occurs. When schools produce comprehensive reports, comprehensive examinations are often administered during a one or two week block of time. Such examinations no doubt encourage students to engage in valuable review activities. When tests are given for formative purposes and results used to make adjustments that maximize learning, extra value is added to the testing process. The results of comprehensive examinations, however, are seldom used in a formative manner. The end-of-term examination, then, may be a less valuable exercise for the student than the formative test. The setting aside of a block of time for testing may be counterproductive, because students are disrupted and deprived of instructional time. As the research on effective teaching indicates, time on task is positively associated with learning gains in the basic skill areas. It makes sense intuitively that the time involved in large-scale formal testing with little follow-up activity may be uncorrelated or even negatively correlated with learning gains. With an approach to evaluation where all available data are used to support a professional judgement, the need for comprehensive testing would likely diminish.

Acquiring test-writing skills. Taking tests properly is a skill that should be taught. Students should be taught how to respond to test items, how to budget their time, and so on. It might be difficult to teach students to respond to constructed-

response items. One method of doing it, however, would be to have classes compose group answers to constructed-response items under the teacher's guidance.

Providing for success. The effects of failure can be very profound, especially if failure is continuous. It is important that students experience success on tests, especially at the beginning of the term. The amount of material covered by a test could be reduced in the early stages of a course and students who do not make a satisfactory grade on a test could be allowed to rewrite a parallel form. However, teachers should be careful not to create in the minds of students the attitude that success in a particular course will come automatically.

Effects of limited time. The purpose of giving tests is to determine the extent to which students have achieved measurable objectives. It is not a function of the school system to spread students out on a continuum of achievement. Therefore, adequate time should be provided for students to complete tests so that time does not become a discriminating factor. Many students like to check their work before they hand in a test, so the test should not take up the whole testing period. For example, allow students forty minutes to write a thirty-minute examination.

Even when adequate time for a test is provided, there is still pressure for many students. It is therefore unfair to require student to do something on a test that is

difficult to do under the pressures of time. For example, a test may not be the most appropriate time for students to write a poem, or to write as many as three essays. Students cannot move from question to question, answering each without having to stop and think. They have to gather and organize their thoughts whether they write an essay on an English test or solve a problem on a mathematics test. Thinking time should be considered when determining the appropriate length for a test.

High achievers. Some teachers may feel that it is unfair to high achievers if they are not given the opportunity to stand above the crowd. However, such an opportunity may be given by including bonus items of greater than average difficulty on a test. Marks gained on these items can be used to cancel out those which were lost through careless mistakes. Thus, high achievers have the opportunity to make a higher score than would ordinarily be the case.

Surprise tests. A good evaluation system has no surprises. Surprise tests have no place in an evaluation system, because students have no chance to thoroughly prepare in order to show teachers what they can do. Students should always know what is to be tested, how it is to be tested, and when it is to be tested. They should also know how test results fit into the overall evaluation plan.

Ambiguous items. The efficacy of a test depends on its ability to provide information on what students know and

can do. This can be determined only if a student knows what a question is asking. Every effort should be made to ensure that tests contain no ambiguities such as those created when new types of questions are included. Teachers may create excellent test items, but if they are of a type that students have had no experience with, confusion may result. Therefore, only item types that students are familiar with should be included on classroom tests.

Testing for a purpose. Tests measure achievement of course objectives, and should not be used to discipline students. Often, marks are deducted when students' writing is poorly organized or standard English is not used. This is not always a defensible practice. If an objective of a course states that students shall write a well-organized essay on some topic or other, the teacher is justified in deducting marks if it is not well-organized. If course objectives make no reference to such requirements, however, it is unfair to deduct marks. If marks are deducted in this case, the feedback students receive through marks is not clear, since a mark reflects achievement on two objectives - the one being tested and the use of standard English. How much of the mark which can be attributed to achievement on either objective is not clear. This does not mean that teachers can ignore the quality of students' writing on tests. On the contrary, when weaknesses in writing ability or language skills show up on tests, or any other work, a strong signal is sent to the teacher that there is a need for the

teacher to teach standard English and require the use of standard English on all written work.

Immediate feedback. Since many tests given throughout the school year would be used in a formative manner, the feedback which they provide to students is very important. For feedback to be effective, however it should be immediate. Although it is practically impossible to provide immediate feedback on test results, it should be given as soon as possible after tests have been written. If teachers find it very difficult to grade papers within a space of two or three days, the format of the test should be questioned. Any objective that can be measured with the use of a selected-response item should not be tested by a constructed-response item. Selected-response items are difficult to develop, but once a teacher has a bank of them, savings in marking time will be tremendous.

Item sampling. Generally, the value of various sections of a test should reflect the time spent teaching. It would be unfair to sample only lightly an area that took most of the teaching time during a unit. Conversely, it would be inappropriate to sample heavily an area that took very little teaching time.

Providing formulae. With respect to mathematical and scientific formulae, it is important that students be able to apply formulae and recognize where they should be used, not that they memorize them. Therefore, such formulae should be provided whenever they are to be used,

unless it is a stated objective of the course that students should memorize them.

Structure for extended-response items. Students need structure to answer questions that require an extended response. The amount of structure may vary from specifying the approximate length to listing the major points to be discussed. Students should also be told beforehand, of the criteria according to which extended-response items are to be evaluated.

Averaging marks. When teachers have to use grades to help determine the final or term standing of students in certain subjects, average grades are often calculated and used. This may be defensible in subject areas composed of discrete units and where subject material is not sequential. It is not defensible, however, to average grades in subject areas such as mathematics where conceptual development and skill acquisition are sequential and cumulative. To average in these subjects penalizes a student for having failed to achieve an objective tested on an exam in October, even though the student may achieve the objective in November.

Using test results. Unless all the objectives of a course, excluding the affective, can be measured through tests, it is unfair to evaluate a course solely on the basis of tests, because test results provide a limited view of student performance. For a comprehensive evaluation, where progress on non-measurable objectives is considered, other techniques should be

used to gather information on their achievement. However, a student who passes a final comprehensive test that carries significant weight should pass the course.

Using publishers' tests. Publishers' tests often do not measure achievement of the objectives of a program as they have been defined in curriculum documents. They do not always do a good job of measuring the objectives stated in the textbooks they accompany. Only those that measure the objectives outlined in curriculum guides and those that have been constructed in accordance with test construction techniques outlined in **Appendix J**, should be authorized.

Reducing test error. All test scores are composed of a true score plus or minus error score. One possible source of error in test scores is surface quality of the test. If students have trouble reading a test that is poorly written or poorly copied, the amount of error in their scores will be larger than would normally be the case. Teachers should ensure that tests are very clearly written and very clearly copied.

Work Sample Files

Work sample files or portfolios containing dated work samples that a student has produced throughout the year provide useful evaluation data. If the samples are systematically collected from all areas of

study regularly, it will provide an overall view of a student's progress.

The work sample file can be used for both formative and summative evaluation, and it is also an excellent aid in talking to parents about a student's strengths and special needs.

In addition to dated work samples, the work sample file could also contain oral reading tapes, anecdotal records, checklists and rating scales. Students should be given the opportunity to select some of the samples to be included.

Advantages of the work sample file, some of which are noted in Cryan (1986), include the following:

- The collecting of samples is not time-consuming.
- Work samples may be used as direct evidence of progress.
- Samples may be interpreted later and by persons other than the teacher.
- Work samples may be used in conjunction with other evaluative assessments.

The work sample file should be an important feature of any language arts writing course but it can also be a feature of any course at any grade level. Wolf (1989) states "when students maintain portfolios of their work, they learn to assess their own progress as learners and teachers gain new views of their accomplishments in teaching" (p.35).

Summary

In this chapter on sources of evaluation data, the more prominent methods of gathering information to ensure appropriate and effective student evaluation have been highlighted. These sources can provide data for pre-instructional, formative and summative evaluation.

While different sources may be more appropriate at primary, elementary, intermediate and secondary levels, teachers should make a concerted effort to use a variety of sources. Student evaluation is not comprehensive and cannot be considered adequate if it is based solely on one source of evaluative data. Furthermore, after a source has been deemed appropriate for a particular situation, it is incumbent on the evaluator to fully understand the advantages and limitations of that source. It should be realized that accurate and valid data can be obtained only if sources are used within the suggested guidelines.

INTERPRETING AND USING EVALUATION DATA

When a student's progress is being summarized for reporting or for the making of an educational decision, all relevant information should enter into the teacher's judgement of that student. In addition to examination results, the data should include the results of observations, work samples, the student's own impression of his/her progress, and any other available information.

Although little has been said about teacher judgement in this handbook, it is a very important ingredient in any evaluation decision. Where teacher judgement is absent, evaluation has not occurred. Informed teacher judgement is necessary in any evaluative exercise, and nothing should impinge upon the right and the responsibility of the teacher to make that judgement. For the most part, evaluation is based on objective data, but teachers may have to go beyond objective data gathered in the classroom. When such factors as a traumatic event in the family is judged to have had an influence on student achievement, they must be taken into account. Thus, teachers must make judgements, interpret evaluation data, and make defensible decisions about grading, reporting and promotion.

Grading

In educational debates, many questions have arisen regarding grades. Simon and Hart (1973) questioned why grading systems were so prevalent when they

might actually serve to divide teachers and students. In a survey of high school teachers, Burton (1983) found that 25% of them believed that grades could affect a student's self-concept. Mazurek (1980) noted that a "B" did not mean much to students unless they knew what an "A", a "C", or an "F" looked like. Holmes (1982) referred to the difficulty of reliably grading English papers. The most detrimental effect of grades occurs when they become ends in themselves for both students and teachers, and when they become more important than learning (Burton, 1983; Indireran, 1983; Zimmerman & Talbot, 1970).

Grading is a necessary and pervasive practice in education in spite of all the weaknesses in its application in the classroom. There are many reasons to defend grading systems. Grading is used to motivate students; to provide a basis for decisions about promotion, graduation, and subsequent schooling; to recommend students and graduates for employment; to award special honours; to provide information for counselling, and to inform parents of the progress their children are making.

Types of Grading Systems

At present many grading systems exist, but the most pervasive systems used for reporting achievement are percentages, letter grades, or pass/fail. Percentage grades and letter grades share many of the same advantages and disadvantages. Both are convenient to use for

for administrative purposes. Both are readily accepted by post-secondary institutions, since they are fairly easy to interpret and they are good predictors of future grades. For some students, percentage and letter grades may be motivators. One advantage that percentage grades have over letter grades is that they are simple and quantitative, and they can be entered as they are into many statistical calculations. This is difficult to do with letter grades, since they usually represent a range of achievement rather than a precise quantity.

Letter and percentage grades also share many of the same disadvantages. They do not give a detailed indication of students' strengths and needs, and they are sometimes based on a number of variables other than achievement. This makes grades difficult or impossible to interpret, and reduces their reliability. The use of percentage or letter grades contributes to anxiety and excessive competition among students, and may inhibit the development of their self-confidence. Grades given in different schools by different teachers may not be comparable, since different standards are often applied in arriving at them. A disadvantage which does not apply as much to letter grades as it does to percentage grades relates to interpretation. The latter are often interpreted as if they were perfectly valid and reliable and without measurement error in the instruments or processes used.

No letter or percentage grades are used in a pass/fail grading system. Pressures of

competing for percentage or number grades are removed, and creativity and exploration of new subject matter is encouraged. This system promotes an atmosphere conducive to learning, although some pressure may be felt by students close to failing. The amount of information provided by pass/fail grades, however, is less than that provided by percentages and letters. Such grades do not discriminate among students of differing abilities, and they are poor predictors of future academic success. Perhaps the most severe limitation of pass/fail grading systems is that they encourage minimum rather than maximum achievement.

Bases for Grading

Grades can be based upon many criteria. Some of the more common ones are the extent to which the student has achieved objectives, the progress which a student makes over some period of time, achievement in relation to ability, and achievement in relation to others in the class. Additionally, some teachers use effort, interest, cooperation, attendance, neatness and diligence as bases for grading, but these are seldom used as complete measures which are independent of actual achievement.

A grading system which uses more than one criteria for grading is unacceptable, since grades lose their meaning when they are based upon a number of criteria. A general rule is that, if more than one

criteria is to be used for determining a grade in a subject, a separate grade for each criteria be reported. If grades are to be based on interest, effort, progress and achievement, there should be a separate grade for each variable. If this tends to make grading too complicated, it is better to omit some grades rather than lump them all together.

Teachers use a variety of procedures in grading students' work on a day-to-day basis. Work may be graded with a letter, a percentage, an anecdotal comment, a rating scale, or any combination of the various schemes. All of these procedures are designed to give students a measure of how well they have performed in relation to stated objectives. It is generally agreed that anecdotal comments are extremely valuable since they can identify clearly those objectives where further instruction is necessary. For students who have achieved the objectives being evaluated, an anecdotal comment can be a source of inspiration and motivation.

When a student's achievement is summarized, relevant information from all data sources should be used. The summative evaluation should be a general assessment of a student's performance, and should not be seen as an opportunity to average the grades which the student obtained from various evaluations. When grades are reported to parents or guardians in *term reports*, teachers may wish to give a general rating using a limited number of categories. For example, at primary, elementary and intermediate levels, a three-point grading scale supplemented by anecdotal comments that identify objectives with

which a student is experiencing difficulty could be used. The grading procedure could be an individual evaluation of a list of objectives in a specific curriculum area or one general evaluation for a number of objectives. This three-point scale could have categories such as: (1) the student has achieved the stated objectives; (2) the student has achieved the stated objectives to a passing level; and, (3) the student has experienced difficulty with the stated objectives, and the minimum level of achievement has not been met.

In cases where a rating scale with a limited number of factors has been used, comments relating performance to course objectives should accompany the report. Statements such as "can do better" or "should try harder" should be avoided. From time to time, variables other than academic achievement might be evaluated,



but these evaluations should be kept separate. If a student is experiencing difficulty, more specific information should be provided.

Reports containing general ratings and supported by anecdotal comments can convey much more information to parents than any of the traditional grading systems. It is doubtful, for example, that the parents of a student who receives a 62, a "C" or a "Pass" in English can tell how well that student can write. While anecdotal comments are somewhat more time consuming to prepare, the extra information they convey is well worth the effort.

Reporting

Reporting refers to the provision of information regarding the performance of students. Information may be quantitative or qualitative or some combination of both. Quantitative information usually takes the form of numbers such as marks, while qualitative information is usually in the form of a written summary of student achievement in relation to specified objectives. Communication may be one-way such as that provided by report cards, or it may be two-way such as that provided in parent/guardian-teacher conferences.

Many schools have designated reporting periods throughout the school year where a report card conveys summative infor-

mation. These reports should have a format readily understood by the parents or guardians, and provide information on the cognitive, affective, and psychomotor development of students. They should emphasize all aspects of students' work (positive and negative) and serve to focus discussion in conferences with teachers which should follow.

Although formal reporting may be sporadic, provision of feedback should be a continuous process. Much information can flow continuously from school to home or from home to school through informal contacts. Telephone calls, notes to parents or guardians, and the sending home of annotated samples of students' work can provide very specific and meaningful information. In addition, the professional competence and personal interest conveyed through school-initiated contacts can build a positive and supportive relationship.

A potential problem exists where subject-teaching is common as would be the case in most high schools. Since teachers often interact with large numbers of students, the student-teacher relationship becomes somewhat impersonal, and no one teacher is personally responsible for monitoring student progress in all subjects. Often, no one is responsible for identifying signs that a student is experiencing difficulty and for initiating appropriate corrective measures. Even in cases where homeroom teachers are assigned to particular classes, these teachers are usually assigned administrative tasks such as keeping attendance

records. Radwanski (1987), in a report to the Ontario Ministry of Education, recommended that every student in a high school be assigned to a monitoring teacher. He stated:

The monitoring teacher should be aware of the student's academic and personal background. All other teachers should provide the monitoring teacher with the student's test and assignment results on an ongoing basis. They should also promptly inform this teacher of any problems with attendance, class performance or attitude, and particularly of any sudden change in any of these areas. (p.111)

The following advice is offered to schools:

- For intermediate and high school students, performance in all domains should be monitored frequently by an assigned teacher, who should assume the primary responsibility for reporting to parents or guardians.
- For elementary students, homeroom teachers should be given the responsibility of fulfilling the monitoring and reporting function.

Regularly scheduled meetings involving teachers and parents or guardians are essential to a good reporting system. An orientation meeting held early in the school year provides an opportunity for parents and guardians to receive an overview of the school's curriculum, methods, and philosophy while promoting



a sense of personal identification with teachers. This is also an opportune time for parents or guardians to meet with the teacher who will monitor their child's performance for the school year. A conference should be held shortly after reports are sent home. Furthermore, it is important that all parents or guardians receive an oral report from the monitoring teacher at least once each term.

The following list of suggestions adapted from Gronlund (1974) should be considered when planning interviews with parents or guardians.

Preparing for the Conference:

- Have a clear grasp of the purpose of the conference.
- Review the student's school records for general background information.
- Assemble a folder of specific information concerning the student's present learning progress.
- Organize the information to be presented.
- Make a tentative list of questions.
- Anticipate questions.
- Provide a comfortable, informal setting free from interruption.

Establishing and Maintaining Rapport during the Conference:

- Create a friendly informal atmosphere.
- Maintain a positive attitude.
- Use language that is understandable.
- Be willing to listen to parents or guardians.

- Be honest and sincere and do not betray confidences.

Sharing Information with Parents/Guardians during the Conference:

- Begin by describing the student's strong points.
- Describe the areas needing improvement in a positive and tactful manner.
- Encourage parents/guardians to participate in the conference.
- Be cautious about giving advice.
- Don't make predictions.

Planning a Course of Action with Parents/Guardians during the Conference:

- Begin the concluding phase of the conference with a brief overall summary.
- Have parents/guardians participate in planning a course of action.
- Review your conference notes with parents/guardians.
- End the conference on a positive note.

Parents or guardian have a right to know, on a continuing basis, what their children are doing in school (goals of instruction), how progress is to be determined (evaluation methods), and how well their children are doing (evaluation results). A good system of communication should provide information on all three elements. In addition, it must be flexible enough to convey a variety of information in language and format which can be readily understood and interpreted by students

and their parents or guardians. Communication with the home should promote a positive image of the student and a 'can do' attitude by the school with respect to improving, or maintaining progress.

It is difficult, if not impossible, to recommend one best system for reporting. As Stake (1986) stated, the teacher must use a variety of reporting methods to convey the results of evaluations. No one approach can adequately describe all aspects of a student's performance. Teachers and principals should concern themselves with interpreting the school's mission to the community. They should establish and maintain formal and informal communication channels with the home where a two-way flow of communication is maintained.

Promotion

When all evaluations for the school year have been completed, teachers and principals often find themselves having to decide if certain students should be promoted or retained. Traditionally, this province has had a very high number of students repeating a grade. In the past 20 years, we have seen individualized instruction, non-gradedness and social promotion. These have been used to prevent retention and the social ills attached to failing.

Two lines of argument exist regarding grade retention. One argues that students

should progress through school with their peers, while the other argues that students who progress to the next grade without having acquired the necessary knowledge and skills are heading for serious trouble. Doyle (1989), in a review of research on grade retention, stated that the practice has been "thoroughly discredited by research" (p.215). He maintains that "all children have a right to be placed in the learning environment that best insures their educational achievement" (p.218), and is adamant that repeating a grade is not an acceptable solution. Holmes and Matthews (1984) found that low achieving students who were promoted generally made greater improvements than students retained in a grade in both the cognitive and affective domains. However, despite the overwhelming evidence documented in educational research, grade retention remains a common practice.

In this Province, there appears to be some relationship between grade retention and dropping out. Pope and Davis (1984) indicated that academic failure was listed as the prime consideration for leaving school by more than 40% of those contacted.

Statistics from the Annual General Returns indicate that 1789 of the Kindergarten to Grade 9 students repeated a grade in 1989-90, 2263 in 1988-89, and 2690 in 1987-88. Assuming that no student has spent more than two years in the same grade, or repeated successive grades, then 6742 different students repeated at least one grade in that three year period. Also, statistics from the last

two years indicate that approximately 68% of the repeaters are males and approximately 65% are at the intermediate level.

In the 1988/89 school year, there were 1572 student withdrawal forms submitted to the Department of Education. Of those reporting the number of years students had spent in each grade, 1354 indicated that students had repeated at least one grade. The other 218 students either did not repeat a grade, or the number of years in each grade was not reported.

Typical research findings on promotion, identified by Doyle (1989), and Miller (1980), include:

- The main reason given for retaining a student is lack of academic achievement, but research shows that retention does not bring significantly greater gains.
- Fear of possible failure does not make students work harder.
- Greater homogeneity of achievement within a grade level is not achieved through retention.
- The more times students are retained, the lower their self-concept.
- Students retained have more adjustment and socialization difficulties.
- Teachers and peers tend to develop unfavourable attitudes towards students who are retained.

- Low grades and retention may bring unfavourable attitudes from the student's family.
- Retention may be justified in the case of a student who has been absent a great deal, or is very immature.

Smith (1987) stated that students who are retained pay with a year of their lives. They understand that their removal from friends and playmates is caused by failure, and this can be very upsetting to them. Hammack (1986) found that holding students back a year or more in elementary school increases the probability of their dropping out.

The major problem occurs when the curriculum is taken as a given, and any student who cannot keep up is considered to be a failure. This might be regarded as blaming the victim. In this circumstance, the teacher has lost sight of the most basic goal of education – the development of each student to the fullest. This does not mean that one student has to learn what another learns, at the same time, in the same way.

The research does indicate that, for some students some of the time, retention can be valuable. However, the problem lies in identifying the appropriate student at the appropriate time. Schuyler (1985) found that long-term achievement results suggest that retention does not help most students. The typical pattern is that even those who benefit in the short run cannot maintain the improved growth rate. Smith (1987) stated that, although some small

percentage of those retained may be helped, the evidence indicates that educators are simply unable to predict accurately which individuals these will be. She further states that retention discriminates against boys, poor students, the relatively young, and the relatively small.

The question to be answered when deciding upon promotion is not whether the student has advanced far enough academically or socially, but how well the student's needs will be met as a result of the decision. Such decisions, which have far-reaching consequences, should not be made by one person, and the interests of the student should be looked at from more than one point of view. The report of the survey done by the Canadian Education Association (1989) indicates that, in the most Canadian jurisdictions, decisions regarding promotion are made by a panel composed usually of the principal, classroom teachers, other teachers, and, often, parents. It is useful to include parents or guardians since they not only bring extra information to the process, but their involvement helps ensure that students do not develop negative attitudes. In fact, students themselves should participate if they are mature enough. How they feel they should be treated is an important consideration of any promotion decision.

Decisions to promote or retain are not simple ones. In the report of the CEA (1989), it was noted that these decisions are based for the most part on academic achievement. This, however, is not the

only factor that should be considered. An In-Grade Retention Scale made up of 21 factors considered important to promotion decisions is contained in **Appendix K**. One advantage of using it is that it ensures that all important factors are taken into account when a decision on promotion is being made.

Very few students should be retained. Diagnostic and remedial help can be given by the regular teacher in cooperation with, and guidance from, the special education teacher(s) in the school. In a spiral curriculum, the previous years work is reviewed before new concepts are developed. This constant repetition and review would seem to obviate the need for student retention in all but the most extreme cases.

Summary

Grades may be based upon a number of criteria such as the achievement of objectives, progress over time, achievement in relation to ability, and achievement in relation to others in the class. The grade which has the most meaning is based on the extent to which students have achieved course objectives. Even those grades contain no information with respect to students' strengths and weaknesses. Information of this sort can be conveyed only if grades are accompanied by qualitative reports.

Research indicates that only a very small percentage of students benefit from

retention. However, it is difficult to identify these students. Before a decision is made to retain a student in a grade, there should be consultation among the teacher, the principal, and the parents or guardians. No student should be retained in a grade if satisfactory accommodation can be made for that student in the next grade.

Two-way communication between the home and school sets the stage for optimal student achievement. This communication should be continuous and of an informal and a formal nature. Though much information can be conveyed to the home through grades referenced to the objectives of instruction, formal reports should be followed up with parent-teacher interviews. When they are properly planned, these interviews provide a rich form of two-way communication which would be difficult to obtain through other means.

CHAPTER V

SPECIAL ISSUES IN STUDENT EVALUATION

In this chapter, a number of issues have been singled out for special attention. These issues deserve more attention than that received in the general discussion on the evaluation of student achievement.

The Public Examination Process

When considering the process of evaluation and its characteristics as presented in this handbook, questions have been asked as to whether public examinations are an acceptable component of a comprehensive, child-centered approach. Are public examinations in conflict with various tenets of this handbook and policy guide? Are adequate provisions made for differentiated evaluation? Do public examinations control the evaluation process rather than complement other activities in that process?

It has been pointed out that course objectives must guide the evaluation process, and this is the case in the preparation of public examinations. There is a concerted effort to ensure that examination items relate to course objectives and span the various levels of cognitive skills. Public examinations make up 50% of the final grade and evaluates only those objectives that can be measured on a valid and reliable examination. Schools, in their share, are responsible for evaluating all the objectives, including those that cannot be reasonably evaluated

using examinations. However, the affective objectives should not be used to determine final grades. Likewise, since public examinations evaluate product, teachers should consider the process component in their portion of the shared evaluation if there is to be a balance of process and product in evaluation. There should be a deliberate attempt by teachers to counterbalance the public examination emphasis on cognitive skills and product, by using other sources of evaluation data in their evaluation.

The Senior High School Certification Handbook points out that where special circumstances dictate, alternative arrangements can be made for evaluation. It is essential that teachers discuss all special cases with their principals, and principals should take the leadership role in securing necessary modifications. This may include a request to defer an initial public examination evaluation, provision of alternate means, some modification of the examination process or modification of the examination itself. Modification of the public examination procedure can be considered for any case deemed exceptional.

There is a general feeling that the public examination structure and the subsequent process of calculating or adjusting final grades is rigid and inflexible. While there are specified criteria, any situation or decision of the Department of Education which is considered detrimental to a student or group of students can be

reviewed. Examinations can be rechecked or reread and initial modifications can be reviewed. Every effort is made to ensure that final results are true indicators of the student performance as measured in relation to stated course objectives.

The modification of school marks is a statistical procedure used to make a fairer system of evaluation, but as with all such procedures, it has a margin of error which can be dealt with through an appeal procedure. There are divergent views on the use of the formula. No procedure will satisfy everyone. There is a segment demanding elimination of the modification procedure and there are others calling for a more rigid approach including the adjustment to 0 rather than allowing for a margin of error of ± 5 . At present, adjustments are calculated and applied on a school basis, which implies that particular classes or individual students will face adjustments because of discrepancies caused by other classes or other students. Obviously, there are no quick solutions to these and other issues, and a more indept study is required before alternative procedures are implemented. However, since the Department of Education recognizes that there is a margin of error in all statistical and testing procedures, all special cases are given full consideration. In that respect, public examinations and subsequent adjustments to derive final marks are truly child-centered.

Language Across the Curriculum

"Language across the curriculum" focuses on the important role of schools in

creating or producing literate citizens. Literacy amongst students should be a high priority and should be a concern in all subject areas. However, the attempt to develop language skills in all subject areas becomes a problem if marks get deducted in all courses for format and language errors.

All teachers have an obligation to be involved in language improvement across the curriculum. Teachers should outline rules for students in assignment writing, specifying the standard of performance required in format, spelling, etc. However, it is felt that teachers should be extremely careful about the way in which marks are deducted for English in all subjects. Furthermore, there is a danger when deducting marks for grammatical errors that students may have marks deducted for the same mistake twice or more. As well, if answers are poorly written, teachers often unconsciously deduct marks because of lack of clarity and poor usage. If there is also a formal deduction for this, students lose again, even to the extent of losing marks for the same error in each question on a test. Consequently, regardless of the subject and objectives being tested, students may lose several marks for one error.

Students should be encouraged to write as well as they can and if problems are identified, assistance should be provided. However, when a question or a problem is assigned in mathematics, science, history, etc., the degree to which students understand and can answer the question or solve the problem should be what is reflected in the mark. Whether the source of evaluation data is oral, written, or

based on observation, students' understanding of the problem is what is most important. This, however, does not absolve teachers from their responsibility for ensuring good language skills across the curriculum, nor does it contradict the view that all teachers have an important role in language development.

The Comparability of Grades

Many questions arise if common grades are assigned to students having different programs and to students evaluated differently on the same program. This is particularly evident if students have individualized program plans (IPP) and if teachers use different sources of evaluation data to evaluate different students on the same course objectives. Each situation should be dealt with on an individual basis, and the assigned grade should be supplemented by anecdotal comments. For the IPP student, specified objectives can be listed and achievement or non-achievement rated. For situations involving differentiated evaluation, special provisions used in the evaluation process can be noted.

It is important that an evaluation represent a student's achievement in relation to stated objectives. For the majority of students, these objectives would be similar. Where there are exceptional cases, the recommended grading system with its limited number of categories and anecdotal comments should provide the necessary flexibility for a comprehensive evaluation. Furthermore, an evaluation is a value judgement of a

student's performance which should consider individual circumstances. Absolute standards for comparison of one student with another cannot be the primary issue.

Teachers should clearly communicate to parents or guardians what a reported evaluation actually means. For the IPP student, as well as the regular student, parents or guardians should have an understanding of how well their child has achieved the stated course or program objectives.

Time Spent on Testing

The issue of testing in education has always been a controversial one. Proponents of educational testing argue that tests are necessary tools for evaluating achievement, and that students cannot be comprehensively evaluated without their use. Furthermore, they argue that test results aid learning when they are put to formative use. Also, tests are seen as helpful to teachers in that they help them clarify the objectives of instruction as they prepare them. Perhaps the most commonly used argument in favour of tests is that they serve as motivators for students.

Critics of testing, on the other hand, have argued that tests psychologically harm children who do not do well on them, since they prevent the development of a healthy self-concept. They maintain that tests create anxiety for most students, categorize and label students, and often create self-fulfilling prophecies where

teachers do not expect much of students who do not do well on tests. Other criticisms have to do with the mechanical use of test results and the exclusion of other useful information from the evaluation process. There are elements of truth in all arguments for and against testing. However, it is fair to say testing has a role in education and is here to stay. Nevertheless, there are legitimate concerns which must be dealt with. One of these relates to the amount of time spent on testing. Many educators and parents argue that an inordinate amount of time is devoted to testing, especially at the intermediate and high school levels. In addition to the time devoted to necessary formative testing in the classroom, large blocks of time are being set aside in many schools for midterm and end-of-year summative testing programs. During this time students miss many hours of instruction.

Researchers such as Denham and Lieberman (1980) point out that the accumulation of academic learning time is a positive correlate of achievement. Their Academic Learning Time (ALT) model states that as students attend to learning tasks and experience a mixture of high and moderate success rates with little or no time spent in low-success activities, they accumulate academic learning time which is positively associated with achievement gains. According to Denham and Lieberman, learning students spend more time working on tasks directly related to the subject matter to be learned. They are very attentive and busy performing the academic part of the task. Where block testing programs are in effect, the opportunity for students to

accumulate academic learning time is reduced since instruction time is reduced. To some extent, large scale block testing can be counterproductive.

Comprehensive examinations, however, serve a useful purpose, since they provide the time for review and the information that is needed to evaluate student performance. Those who support less testing time argue that evaluation data, if gathered carefully over time, provides sufficient information on which to base grades. Furthermore, they consider data gathered through chapter and unit tests, observation, assignments, projects, and portfolios to be more valid and reliable than a grade which is derived in large part from a single, comprehensive examination.

The discontinuation of block examinations in many subjects could have a number of advantages. First, the amount of available instructional time would increase, creating the possibility that the accumulation of academic learning would increase. If academic learning time is increased, research tells us that achievement will probably increase. Second, student evaluation practices could improve. Greater care would have to be exercised for the collection of all data that will be used in producing a final grade.

CHAPTER VI

SAMPLE POLICY

This chapter contains sample policy statements which are based on positions outlined in this handbook. Each section begins with a broad policy statement. This statement is followed by a rationale which is a synopsis of the section upon which it is based as well as general guidelines which support the policy.

The handbook and policy statements could be used as the basis for developing a district policy. It might not be appropriate to adopt the policy statements as they are, since those affected by them will not have been involved in their development. Teachers and principals should have some ownership of any policy that affects their work. This is especially true in such a critical area as student evaluation.

At the district level, there may be a desire to tailor the handbook to meet local needs. Certain issues may be expanded, or there may be a need to address some issues which have not been raised. In any event, teachers and principals should be involved in the process of making this or any other policy document fit the needs of a particular district. More effective implementation will likely be the result of this involvement.

Evaluation of Student Achievement

Policy Statement: Student evaluation practices will be based on a philosophy of education which respects the uniqueness of each child and be conducted according to currently modern educational theory and practice.

Rationale

Evaluation refers to a professional judgement of the extent to which students have achieved objectives. It is not a discrete activity, but a continuous and comprehensive process, the results of which always affect classroom proceedings.

Objectives represent goals for students, and they provide the framework for student evaluation. These objectives are concerned with skills and values as well as with knowledge. They are not always measurable, or even observable; they do not always specify a product, and their value may be intrinsic.

There are many procedures which educators may use in evaluating students. None of these procedures is infallible, and their suitability often depends on the nature of the subject or discipline in which the evaluation occurs. A primary consideration in this instance would be the validity and reliability of the instruments used. Ultimately, the evaluation procedures to be used are determined by objectives.

Parents are important partners in education. They deserve to be fully aware of the progress their children are making at school. Not only should parents be made aware of their children's progress, but their support for the efforts of the school should be encouraged.

Finally, evaluation must always be a purposeful activity.

Types of Evaluation

Policy Statement: *Evaluation will consist of pre-instructional, formative and summative activities.*

Rationale

Comprehensive evaluation requires an awareness of the strengths and weaknesses that students bring to the classroom. This information may be gathered from such sources as cumulative records, teacher-teacher conferences, and parent-teacher conferences before instruction begins. This type of evaluation, referred to as pre-instructional, should ensure the selection of appropriate instructional strategies.

The most valuable type of evaluation occurs during the instructional process. Formative evaluation focuses upon the process as well as the products of learning. Its main function is to document progress and identify impediments to learning such that they can be corrected or remediated as soon as possible. Although the sources of formative evaluation data are varied, a very important one is teacher observation.

Summative evaluation is used to grade or certify students after instruction has taken place. A very common source of summative evaluation data is the teacher-made test. These tests, if properly constructed, can provide measurements which, when used in conjunction with evaluation data gathered through other sources, can provide a comprehensive picture of student achievement and progress.

Guidelines

1. Use available sources to determine the strengths and weaknesses of students.
2. Emphasize formative evaluation during instruction.
3. Gather information for formative evaluation by using a variety of data sources.
4. Communicate formative evaluation data promptly to student, parents, and the student's homeroom (or monitoring) teacher.
5. Summative evaluation decisions should be based on data from a variety of sources.

Objectives

Policy Statement: *Evaluation represents performance in relation to stated objectives from the affective, cognitive and psychomotor domains. Objectives must be clearly stated and communicated to students.*

Rationale

The Aims of Public Education for Newfoundland and Labrador create a vision of the type of individual who is truly educated. The achievement of these aims cannot be determined until well after schooling is finished. These aims give rise to more narrow goals which are the outcomes of schooling in general rather than at a specific level.

The goals and aims of education are achieved largely through curriculum. Curriculum objectives specify outcomes which are achieved over a significant portion of a student's school life. Curriculum objectives are achieved through course objectives. These, in turn, are achieved through unit and lesson objectives.

Objectives guide teaching and learning. If objectives are to guide learning, students must be aware of them. They should specify what students will be able to do following instruction, and they should be understood by students.

Education is concerned with much more than student recall of knowledge. As students mature, objectives should be written to ensure that students engage in complex cognitive processes where appropriate. They should be required to manipulate knowledge by analyzing it, applying it to solve problems, drawing on different parts to form a new whole, and evaluating it.

The affective development of students is of direct concern for educators. Evaluating in the affective domain, however, is difficult. Specific objectives do not play the same role in that the provision of such objectives cues students to act in a particular manner. The observed indicators, given the presence of such cues, would not necessarily mean that an objective has been achieved. Nevertheless, a clear idea of the affective objectives to be achieved and specific indicators of their achievement is needed. Affective objectives should focus on the student's willingness to pay attention, to take part, to attach value to things, and to begin to build a value system which gradually becomes pervasive, consistent and predictable.

Once affective objectives and specific indicators of achievement have been developed, students can be evaluated through observation with the use of rating scales, checklists and

anecdotal records. They may be evaluated through self-reporting devices such as attitude scales, or individual inventories of attitudes and preferences, through student reports or term papers which require the student to take a position, or through teacher/student conferences where indirect questions are used to ascertain students' attitudes and interests.

It is important that numerical values not be assigned to signify degree of affective development. It is important as well that progress, or lack of it, not affect progress in the other domains. Finally, it is important that parents be given information on the affective development of their children. As the child's first teachers, parents are in a position to positively influence the affective development of their children and play a role which is complementary to that of the school.

The psychomotor domain is concerned with the motor development of students: learning to print or write, using laboratory equipment in science, baking a cake in home economics, bouncing a ball in physical education, and so on. Objectives in this domain should be concerned with the use of the senses, readiness to act, the early stages of learning a new movement, proficient movement, skillful performance, adaption of physical skills, and creation of new movements. Observation with the use of rating scales, checklists, and anecdotal records can be used to evaluate psychomotor development. In addition, measurements can be used in some subject areas.

Guidelines

1. State objectives in language that students can understand.
2. Provide objectives to students at the beginning of a unit.
3. Evaluate students according to the stated objectives.
4. Keep achievement in each domain independent of achievement in the other domains.
5. Where appropriate, write objectives which require different cognitive processes.
6. Use a variety of data sources to evaluate the cognitive domain.
7. Compile affective objectives collaboratively at the school level.
8. Evaluate the achievement of affective objectives by using a variety of sources including observation, student reports, and personal discussions.
9. Do not use achievement in the affective domain as a basis for promotion.
10. Provide objectives for the psychomotor domain for those courses where psychomotor development is important.
11. Evaluate achievement of psychomotor objectives through observation. Measurements may be used where they are appropriate such as in physical education.

Process and Product in Instruction and Evaluation

Policy Statement: Process and product objectives will be evaluated.

Rationale

Product refers to a body of information that the student interacts with and is expected to know. Knowledge is transmitted to students who, at some point, must demonstrate mastery of it. In this product approach to education, the ends are all-important, and any means which lead to them are justified. Where there is a product orientation, students often interact passively with content.

In a process approach to education, the emphasis shifts from passing on information to the methods by which it is acquired. Students are more active as they learn how to find information, how to become autonomous thinkers, how to use what they know to solve new problems, and how to make decisions for themselves. Some advantages of a process approach to instruction are that it increases motivation and intellectual potency, aids memory, teaches students to discover for themselves, increases self-confidence, and improves attitudes toward learning.

Orientations to teaching are not pure process or pure product. There is some concern with both. However, when product becomes an end in itself, the balance between process and product is upset, and process is a slighted partner. Likewise, in a situation where process is all-important, the product essential to the students' full development is ignored. Ideally, the curriculum should be regarded as being worthwhile and of intrinsic value, but the importance of the processes instrumental to the achievement of the product objectives should not be overlooked.

When a balance has been struck between product and process, the extent to which students know and comprehend things, and the extent to which they can do such things as think autonomously, use prior knowledge to solve new problems and to make decisions, should be evaluated. Also, the evaluation should consider development in the various cognitive skills that students need to have to be able to do these things.

The full range of data sources may have to be employed to provide meaningful evaluations in all areas.

Guidelines

1. Reflect the relative emphasis given to product and process during evaluation.
2. Use many different data sources in the evaluation of process.
3. Provide students with general evaluation plans at the beginning of a course.

Differentiated Evaluation

Policy Statement: *Differentiated evaluation will be employed to accommodate students with special needs.*

Rationale

Education is a basic right of all children in our society. The Aims of Public Education challenge all educators to develop in each individual not just a set of skills, but a sense of personal dignity and worth.

In the process of evaluating students, the evaluation method should not negatively affect their ability to demonstrate achievement of objectives. Objectives may have to be evaluated differently for some students if they are to demonstrate their level of achievement accurately. A student may be unable to demonstrate acquired knowledge because of some special need. However, this does not always indicate that there is a lack of learning or understanding. A different mode of evaluation, may produce completely different results. There is an obligation to focus on the individual needs of a student, and to employ evaluation methods that clearly determine to what extent the objectives of a course have been achieved.

Each student is unique. The range of variation defies description. The school should provide for individual differences among students. Specified objectives and evaluation methods may have to be adapted to meet the needs of students, and where adaptations are made, parents and students should be informed in a timely, sensitive, and meaningful fashion.

Guidelines

1. Reflect the objectives of instruction in evaluation methods.
2. Ensure that data sources are congruent with the needs of students.
3. Explain evaluation modifications to students and parents prior to instruction.
4. Involve teachers in all decisions regarding differentiated evaluation.

Sources of Evaluation Data

Policy Statement: *A variety of sources will be used in the gathering of evaluation data.*

Rationale

A single data source gives a limited view of student achievement. A comprehensive picture of where students stand in relation to course objectives can be obtained only through the use of a variety of data sources.

Observation is a valuable source of data for evaluating process as well as product. To be effective, it must be conducted purposefully, regularly and systematically. Appropriate methods, such as the completion of a checklist, rating scale or anecdotal records should be used to document the results of observation activities. The use of such sources as the work sample folder, self-evaluation activities, interviews and learning journals can also be used to convey timely, relevant information to parents or guardians and students.

Projects are often assigned to students. Generally speaking, however, the younger the student, the more teacher guidance and assistance required. Projects that are assigned at the primary and elementary levels should generally be completed at school, and should be regarded as learning activities, not evaluation activities. At the high school level, the results of assigned projects may be used in evaluating students. Nevertheless, even very mature students need guidance when projects are assigned. Students should be provided with the objectives of the project, a description of the task, relevant dates, references, and evaluation criteria.

Homework assignments may sometimes provide information for evaluation decisions, although their evaluation is restricted to product only. The results of homework assignments can provide valid information if students are properly prepared beforehand. Except in the case of very mature or independent students, any instruction which has to be provided in order for students to complete homework assignments should be provided before the homework is assigned.

Tests should be designed to measure student achievement of measurable objectives. Students should always know what is to be tested, how it is to be tested, when it is to be tested, and how test results fit into the general evaluation plan. Test items should be

written according to established guidelines. They should be completely free from ambiguities, clearly produced, and clearly copied. They should clearly convey to students the type of response they are to give. Time should not be a discriminating factor, and thinking time should be allowed for when tests are constructed. When students experience failure in testing situations, some consideration should be given to providing an extra chance for success, such as an opportunity to write a parallel form.

Although test results may carry considerable weight when evaluation decisions are made, these decisions should not be made on the basis of test results only. Averaging of test results does not provide a valid picture of student achievement in subject areas where conceptual development and skill acquisition are sequential. Often, extra information will be needed to make a decision that is in the best interest of the student since testing cannot be regarded as comprehensive evaluation. Results of many tests, particularly those having selected-response items, provide information only on product. Testing is, therefore, but a subset in the domain of evaluation.

Guidelines

1. Use more than one data source in evaluating student progress.
2. Conduct observations regularly and systematically.
3. Document the results of observations through the use of checklists, rating scales and anecdotal records.
4. Allow primary and elementary students to complete projects at school where teachers can provide the necessary assistance and guidance.
5. When assigning projects, provide students with a written guide containing the objectives of the project, a description of the task, relevant dates, references, and evaluation criteria.
6. Except in the case of very mature or very independent students, provide instruction to students before homework is assigned.
7. Use evaluation data gathered from homework for formative evaluation.
8. Course objectives emphasized during instruction should receive emphasis on tests.
9. Give proper attention to test validity and reliability. Use of publishers' tests should be restricted to test that are rigorously constructed and measure the objectives outlined in curriculum guides.

10. Ensure that students know what is to be tested, how and when it is to be tested, and how test results fit into the general evaluation plan.
11. Do not use tests as a control mechanism.
12. Teach students how to write tests.
13. Construct test items according to established guidelines.
14. Ensure that tests are clearly produced, clearly copied, and contain no unfamiliar item types. Clear up any real or perceived difficulties.
15. Select test items on the basis of how well they measure course objectives, not on the basis of maximum discrimination. Discriminate at the higher levels of achievement in such a way that other students are not penalized.
16. When testing, use items requiring different cognitive processes.
17. Allow more time than would be necessary for the average student to complete a test. Do not require students to do something which would be unreasonable under the constraints of a time limit. Allow for thinking time when constructing tests.
18. Do not average formative evaluation results to obtain a summative result.
19. Allow a student who passes a final, comprehensive test in a course that is amenable to measurement to pass that course.
20. Do not conduct block mid-term testing.
21. Do not conduct block end-of-year testing up to and including Grade 9.

Grading

Policy Statement: *For summative evaluation, grades will indicate performance in relation to the stated objectives.*

Rationale

Summative evaluation is an attempt by a teacher to give an assessment of a student's performance after considering all relevant data, which have been gathered from a variety of sources. It could consist of a rating based on an observation, a letter grade from a project, and a percentage grade on a test. The challenge is to integrate these individual assessments to provide a grade in relation to course objectives. In some cases, the evaluation may be grading of individual objectives; in other cases it may be based on a range of objectives. When more than one type of objective, such as cognitive and affective is being evaluated, a separate grade should be used for each.

When teachers provide a summative grade for a term report, it should be realized that there is still a formative dimension. Teachers, as well as parents and students, are expected to use the information for future decisions about the teaching-learning process. The provision of anecdotal comments can identify a student's strengths, and pinpoint special needs that may require remedial attention. However, there is still a need to use some grading scheme, ranging from broad categories at the primary level to specific percentage grades for senior high courses.

Guidelines

1. Base grades on the extent to which students have achieved the stated objectives of a course or program.
2. The same grades given by different teachers using the same evaluation policy should have, at least approximately, the same meaning.
3. For primary and elementary grades, use a grading system with a minimal number of categories for both term and year-end reports. Grading should apply to the general program objectives, and anecdotal comments should be an integral part of the grading scheme.

4. In intermediate grades, a grading scheme using a minimal number of categories may be appropriate for term reports, but a more specific scale such as letter grades or percentages could be used for year-end grades. Anecdotal comments should be used to supplement the grade, especially on the term reports.
5. For senior high courses, a more specific letter grade, percentage grade, or some modification of the specific percentage grade should be used for term reports. For the year-end summative grade, percentages would be used. Anecdotal comments should supplement the term reports.
6. When percentage grades are used, a predetermined formula should not be used to calculate results.

Reporting

Policy Statement: *A system of formal and informal communication will be maintained in each school.*

Rationale

Parents have a right to know how their children are doing at school. If they are kept informed, then they are in a position to positively influence their children's achievement. Communication can be qualitative or quantitative; it can be one-way or two-way. Quantitative reports of student achievement provide numerical grades; qualitative reports provide statements about achievement of specific objectives. One-way communication occurs when schools send home reports; two-way communication is most often established through teacher-parent/guardian conferences. Most information can be provided in a qualitative form. This is best provided through two-way communications such as that established in teacher-parent/guardian conferences.

Formal reporting is often infrequent in that it takes place only during designated times of the year. Informal networks should be set up so that information flows continuously between the home and the school. Much information can be exchanged through telephone calls, the sending home of annotated work samples, and so on. The best method to use is that which best serves the purpose of continuously providing the home with relevant information.

In most cases, students are exposed to more than one teacher, and no one is responsible for monitoring progress. Therefore, each student should be assigned to a monitoring teacher who would receive evaluation data from other teachers as soon as it becomes available. This teacher uses these data to provide timely feedback during informal contacts with the home.

Guidelines

1. Communicate the results of formative evaluation to students within two or three instructional days.
2. Provide parents or guardians with a report of their children's achievement at least three times a year.

3. Provide for teacher-parent/guardian conferences in conjunction with report cards.
4. Between formal reporting periods, communicate with the home informally.
5. In primary, elementary and intermediate schools, have homeroom teachers monitor the performance of their students. In high schools, each student should be assigned to a monitoring teacher.

Promotion

Policy Statement: *Students will be promoted unless appropriate arrangements cannot be made in the next grade and there is general agreement that promotion is not in the students' best interest.*

Rationale

Retention in a grade is beneficial for only a very small percentage of students. The identification of these students is no easy task. A decision to retain should be made only when no satisfactory accommodation can be made for the student in the next grade. Even then, it should be made in consultation with the student's parents. As a rule of thumb, no student should be retained for failing to satisfactorily achieve the objectives of a course in which learning is not sequential and cumulative.

Guidelines

1. Retain students only when satisfactory accommodation cannot be made in the next grade. When a student is retained, a modified program should be put in place.
2. Before a decision is made to retain a student in the same grade, there should be a consultation process involving the school administration, teachers, and the parents or guardians of the student.
3. Retain students only on the basis of achievement in courses in which learning is sequential and cumulative.

REFERENCE LIST

- Agnew, E. (1985). *The grading policies and practices of high school teachers*. State University of New York, Albany. (ERIC Document Reproduction Service No. ED 259 022)
- Anderson, L., Evertson, C., & Brophy, J. (1982). *Principles of small-group instruction in elementary reading* (Occasional Paper No. 58). East Lansing: Michigan State University, Institute for Research on Teaching. (ERIC Document Reproduction Service No. ED 223 981)
- Andrews, B., Gardner, J., & Hubbard, K. (1986). *Student evaluation...the bottom line*. Toronto: Ontario Secondary School Teachers' Federation.
- Arasian, P. (1985). Educational testing and public policy: Growth and implications. In R. Stiggins, N. Bridgeford, F. Conklin, & C. Brody (Ed.), *A feeling for the student: An analysis of the art of classroom assessment*. A final report to the U.S. Department of Education, Northwest Regional Educational Laboratory, Portland, OR. (ERIC Document Reproduction Service No. ED 266 165)
- Arnold, J. (1985). *Towards a four-fold process of middleschool classroom evaluation*. A paper presented at the annual meeting of the National Middleschool Association, Baltimore, MD. (ERIC Document Reproduction Service No. ED 274 053)
- Bailey, W. (1973). *Developing performance evaluation systems*. A paper presented at the annual meeting of the Association for Supervision and Curriculum Development, Minneapolis, MN. (ERIC Document Reproduction Service No. ED 117 814)
- Baskwill, J., & Whitman, P. (1988). *Evaluation: Whole language, whole child*. Richmond Hill, ON: Scholastic.
- Bloom, B. (Ed.). (1956). *Taxonomy of educational objectives: Handbook I, the cognitive domain*. New York: Longmans, Green.
- Board of Education for the City of Etobicoke. (1987). *Making the grade: Evaluating student progress*. Scarborough, ON: Prentice Hall.
- Board of Education for the City of North York. (1983). *Look hear developing programs for primary children*. Toronto, ON.

- Brookover, W., Beady, D., Flood, P., Schweitzer, J., & Wisenbaker, J. (1979). *School social system and student achievement: Schools can make a difference*. New York: Praeger.
- Brookover, W., Beamer, L., Eftim, H., Hathaway, D., Lezotte, L., Miller, S., Passalacqua, J., & Tornatzky, L. (1982). *Creating effective schools*. Florida: Learning Publications.
- Brophy, J. (1981). Teacher praise: A functional analysis. *Review of Educational Research*, 51(1), 5-32.
- Brophy, J., & Evertson, C. (1976). *Learning from teaching: A developmental perspective*. Boston: Allyn & Bacon.
- Brophy, J., & Good, T. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp.328-375). New York: Macmillan.
- Bruce, K. (1988). Making the grade or marking time? *Phi Delta Kappan*, 69(5), 383-384.
- Burton, F. (1983). *A study of the letter grade system and its effect on the curriculum*. A paper presented at the meeting of the American Educational Research Association, Montreal, Quebec. (ERIC Document Reproduction Service No. ED 238 143)
- Canadian Education Association. (1989, March). Grade promotion and retention: Practices in Canadian School boards. *CEA Information Note*, Toronto, ON.
- Clark, C., Gage, N., Marx, R., Peterson, P., Stayrook, N., & Winnie, P. (1979). A factorial experiment on teacher structuring, soliciting, and reacting. *Journal of Educational Psychology*, 71(4), 534-552.
- Clark, T., & McCarthy, D. (1983). School improvement in New York City: The evolution of a project. *Educational Researcher*, 12(4), 17-24.
- Coker, H., Medley, D., & Soar, R. (1980). How valid are expert opinions about effective teaching? *Phi Delta Kappan*, 62(2), 131-134, 149.
- Cohen, M. (1983). Instructional, management and social conditions in effective schools. In A. Webb & L. Webb (Eds.), *School finance and school improvement: Linkages in the 1980's*. Fourth annual yearbook of the American Educational Finance Association. Cambridge, MA: Ballinger.

- Cooper, H. (1939). Synthesis of research on homework. *Educational leadership*, 47(3), 85-91.
- Crocker, R., & Riggs, F. (1979). *Improving the quality of education: Challenge and opportunity*. St. John's, NF: Government of Newfoundland & Labrador.
- Cronbach, L. (1971). Test validity. In R. Thorndike (Ed.), *Educational measurement* (2nd. ed.)(pp.443-507). Washington, DC: American Council on Education.
- Cryan, J. (1986). Evaluation: Plague or promise? *Childhood Education*, 62(5), 344-50.
- Denham, C., & Lieberman, A. (Eds.). (1980). *Time to learn*. Washington, DC: National Institute of Education.
- Department of Education. (1981). *Design for social studies*. St. John's, NF: Division of Curriculum and Instruction.
- Department of Education. (1984). *Aims of public education for Newfoundland and Labrador*. St. John's, NF: Government of Newfoundland and Labrador.
- Department of Education. (1987). *Primary curriculum guide*. St. John's, NF: Division of Curriculum and Instruction.
- Department of Education. (1988a). *Primary mathematics curriculum guide*. St. John's, NF: Division of Curriculum and Instruction.
- Department of Education. (1988b). *Special education policy manual*. St. John's, NF: Division of Special Services.
- Department of Education. (1989). *Math match: Fitting the pieces together*. St. John's, NF: Division of Special Services.
- Dewey, J. (1966). *Democracy and education*. New York: The Free Press.
- Doyle, R. (1989). The resistance of conventional wisdom to research evidence: The case of retention in grade. *Phi Delta Kappan*, 71(3), 215-220.
- Dressel, P. (1968). Measurement and evaluation of instructional objectives. In N. Gronlund (Ed.), *Reading in measurement and evaluation* (pp. 11-17). New York: Macmillan.
- Dunkin, M. (1978). Student characteristics, classroom processes, and student achievement. *Journal of Educational Psychology*, 70, 998-1009.

- Ebel, R. (1972). *Essentials of educational measurement*. Englewood Cliffs, NJ : Prentice-Hall.
- Eisner, E. (1985). *The art of educational evaluation*. Philadelphia: Falma, Taylor & Francis.
- Eiss, A., & Harbeck, M. (1969). *Behavioral objectives in the affective domain*. Washington, DC: National Science Supervisors Association.
- Emmer, E., Evertson C., & Brophy, J. (1979). Stability of teacher effects in junior high classrooms. *American Educational Research Journal*, 16(1), 71-75.
- Foyle, H., & Bailey, G. (1986). Homework: Its real purpose. *The Clearing House*, 60(4), 187-188.
- Frost, J. (1989). Assessing programme and instruction in early primary grades. *The Canadian School Executive*, 8(7), 11-15.
- Gagne, R. (1974). Educational technology and the learning process. *Educational Researcher*, 3(1), 3-8.
- Gall, M., Fielding, G., Schalock, D., Charters, W., & Wiczinski, J. (1984). *Involving the principal in teachers' staff development: Effects on the quality of mathematics instruction in elementary schools*. Eugene, OR: Center for Educational Policy and Management. (ERIC Document Reproduction Service No. ED 251 929)
- Good, T., & Brophy, J. (1978). *Looking in classrooms*. New York: Harper & Row.
- Good, T., & Grouws, D. (1975). *Process-product relationships in fourth grade mathematics classrooms*. National Institute of Education Grant NEG-G-00-3-0123). Columbia: University of Missouri, College of Education.
- Good, T., & Grouws, D. (1977). Teaching effects: A process-product study in fourth grade mathematics classrooms. *Journal of Teacher Education*, 28(3), 49-53.
- Griffiths, A. (1987). *The evaluation of scientific processes*. Toronto: Holt, Rinehart & Winston.
- Gronlund, N. (Ed.). (1968). *Reading in measurement and evaluation*. New York: Macmillan.
- Gronlund, N. (1970). *Stating objectives for classroom instruction*. New York: Macmillan.

- Gronlund, N. (1974). *Improving marking and reporting in classroom instruction*. New York: Macmillan.
- Gronlund, N. (1985). *Measurement and evaluation in teaching* (5th ed.). New York: Macmillan.
- Hammack, F. (1986). Large school systems' dropout reports: An analysis of definitions, procedures and findings. In G. Natriello (Ed.), *School dropouts: Patterns and policies* (pp. 20-37). New York: Teachers' College Press.
- Higgins, J., Chnupa, P., Marginet, S., Nordhaus, P., Whitney, D., Berger, J., Dowers, J., Thomas, J., & Rowe, L. (1982). *Using measurement and evaluation to promote learning*. Terre Haute, Indiana State University. Curriculum Research and Development Center. (ERIC Document Reproduction Service No. ED 220 511)
- Holmes, M. (1982). *What every teacher and parent should know about student evaluation*. Toronto: OISE Press.
- Holmes, T., & Matthews, K. (1984). The effects of non-promotion on elementary and junior high pupils: A meta-analysis. *Review of Educational Research*, 54, 225-236.
- Hughes, D. (1973). An experimental investigation of the effects of pupil responding and teacher reacting on pupil achievement. *American Educational Research Journal*, 10(1), 21-37.
- Hutchison, L., & Wonacott, M. (1984). *Assess the progress of exceptional students*. Columbus, Ohio State University: National Center for Research on Vocational Education. (ERIC Document Reproduction Service No. ED 241 728)
- Indiresan, J. (1983). Changing patterns in the evaluation of student performance in India. *Higher Education in Europe*, 18(2), 58-64.
- Krathwohl, D., Bloom, B., & Masia, B. (1964). *Taxonomy of educational objectives: Handbook II, the affective domain*. New York: David McKay.
- Kubiszyn, T., & Borich, G. (1984). *Educational testing and measurement*. Illinois: Scott, Foresman.
- Lavi, Z. (1982). Toward an alternative evaluation policy for kibbutz schools. *Interchange on Educational Policy*, 13(1), 92-96.

- Lightfoot, S. (1978). *Worlds apart: Relationships between families and schools*. New York: Basic Books.
- Manasse, A. (1982). Effective principals: Effective at what? *Principal*, 61(4), 10-15.
- Mazurek, K. (1980). The dreaded task of marking. *Alberta English*, 19(2), 20-23.
- McCall, I., Bryce, T., & Robertson, I. (1983). Diagnostic assessment: Assessing foundation science practical skills in the classroom. *Programmed Learning and Educational Technology*, 20(1), 11-17.
- McCormack-Larkin, M., & Kritek, W. (1983). Milwaukee's project RISE. *Educational Leadership*, 40(3), 16-21.
- McDonald, F., & Elias, P. (1976). *The effects of teaching performance on pupil learning. Beginning teacher evaluation study: Phase II, 1974-76*. Princeton, NJ: Educational Testing Service.
- Metzger, D. (1985). Process versus content: The lost illusion. *The social studies*, 76(3), 115-119.
- Miller, M., Frazier, C., & Richey, D. (1980). Student non-promotion and teacher attitude. *Contemporary Education*, 51(3), 155-157.
- Mirkin, P., & David, S. (1979). *Formative evaluation in the classroom: An approach to improving instruction*. Minneapolis: Minnesota University. Institute for Research on Learning Disabilities.
- Muir, S., & Wells, C. (1983). Informal evaluation. *The Social Studies*, 74(3), 95-99.
- Napier, S. (1976). Grading and young children. In S. Simon & J. Bellanca (Eds.), *Degrading the grading myths: A primer of alternatives to grades and marks* (pp. 232-27). Washington, DC: Association of Supervision and Curriculum Development.
- Parker, J., & Rubin L. (Eds). (1968). *Process as content: Curriculum design and the application of knowledge*. Chicago: Rand McNally.
- Phinney, J. (1982). Observing children. *Young Children*, 37(5), 16-24.
- Pope, T., & Davis, H. (1984). *Leaving early - A study of student retention in Newfoundland and Labrador*. St. John's, NF.

- Popham, W. (Ed.). (1974). *Evaluation in education: Current applications*. Berkely, CA: McCutchan.
- Posher, G., & Rudnitsky, A. (1982). *A guide to curriculum development for teachers*. New York: Longman.
- Pratt, D. (1980). *Curriculum: Design and development*. Toronto: Harcourt, Brace & Jovanovich.
- Purkey, S., & Smith, M. (1983). Effective schools: A review. *The Elementary School Journal*, 83(4), 427-452.
- Quebec Ministry of Education. (1981). *Basic concepts in educational evaluation*. Quebec: Government of Quebec.
- Radwanski, G. (1987). *Ontario study of relevance of education and the issue of dropouts*. Toronto, ON.
- Roid, G., & Haladyna, T. (1982). *A technology for test item writing*. New York: Academic Press.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart & Winston.
- Rowntree, D. (1987). *Assessing students: How shall we know them?* (2nd ed.). New York: Nichols.
- Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979). *Fifteen thousand hours: Secondary schools and their effects on children*. Cambridge, MA: Harvard University Press.
- Sadler, R. (1983). Evaluation and improvement of academic learning. *Journal of Higher Education*, 54(1), 60-79.
- Schuyler, N. (1985). *Does retention help? Perspectives after three years*. Austin, TX: Office of Research and Evaluation. (ERIC Document Reproduction Service No. ED 263 143)
- Schuncke, G. (1981). The uses and misuses of evaluation. *The Clearing House*, 54(5), 219-222.
- Shapiro, P., Lukesavitch, A., & Shapiro, B. (1986). Student evaluation: A change of focus needed. *Education Canada*, 26(2), 26-29.

- Simon, S., & Bellanca, J. (Eds.). (1976). *Degrading the grading myths: A primer of alternatives to grades and marks*. Washington, DC: Association for Supervision and Curriculum Development.
- Simon, S., & Hart, L. (1973). Grades and marks: Some commonly asked questions. *The Science Teacher*, 40(6), 46-48.
- Simpson, E. (1972). *The classification of educational objectives in the psychomotor domain: Vol. 3, the psychomotor domain*. Washington, DC: Gryphon House.
- Smith, M., & Shepard, L. (1987). What doesn't work: Explaining policies of retention in the early grades. *Phi Delta Kappan*, 69(2), 129-134.
- Stake, R. (1986). Grading, reporting, and communicating student achievement. In B. Andrews, J. Gardner, & K. Hubbard (Eds.), *Student evaluation...the bottom line*. Toronto: Ontario Secondary School Teachers' Federation.
- Stallings, D., & Mohlman, G. (1981). *School policy, leadership style, teacher change, and student behavior in eight high schools*. Washington, DC: National Institute of Education. ERIC Document Reproduction Service No. ED 209 759)
- Stallings, J. (1980). Allocated academic learning revisited, or beyond time on task. *Educational Researcher*, 9(11), 11-16.
- Stallings, J., Cory, R., Fairweather, J., & Needels, M. (1977). *Early childhood education classroom evaluation*. Menlo Park, CA: Stanford Research Institute.
- Stallings, J., Cory, R., Fairweather, J., & Needels, M. (1978). *A study of basic reading skills taught in secondary schools*. Menlo Park, CA: Stanford Research Institute.
- Stallings, J., & Kaskowitz, D. (1974). *Follow-through classroom observation evaluation, 1972-73*. Menlo Park, CA: Stanford Research Institute.
- Tenbrink, T. (1974). *Evaluation: A practical guide for teachers*. New York: McGraw-Hill.
- Thomas, W. (1986). Grading - Why are school policies necessary? What are the issues? *NASSP Bulletin*, 70, 23-26.

- Tikunoff, W., Berliner, D., & Rist, R. (1975). *An ethnographic study of the forty classrooms of the Beginning Teacher Evaluation Study known sample* (Technical Report No. 75-10-5). San Francisco: Far West Laboratory. (ERIC Document Reproduction Service No. ED 150 110)
- Tomlinson, J. (1981). The troubled years: An interpretative analysis of public schooling since 1950. *Phi Delta Kappan*, 62(5), 373-376.
- Tuckman, B. (1979). *Evaluating instructional programs*. Boston: Allyn & Bacon.
- Weber, G. (1971). *Inner-city children can be taught to read: Four successful schools*. Washington, DC: Council for Basic Education.
- Willis, G. (1978). *Qualitative evaluation*. Berkeley, CA: McCutchan.
- Wittrock, M. (1986). *Handbook for research on teaching* (3rd ed.). New York: Macmillan.
- Wolf, D. (1989). Portfolio assessment: sampling student work. *Educational Leadership*, 46(7), 35-47.
- Wolf, R. (1979). *Evaluation in education*. New York: Praeger.
- Zais, R. (1976). *Curriculum: Principles and foundations*. New York: T. Crowell.
- Zimmerman, R. & Talbot, T. (1970). *Individualization of instruction: Evaluation a roadblock*. (Education Monograph No. 5), Missoula: Montana University. (ERIC Document Reproduction Service No. ED 083 163)

APPENDIX A: LEVELS OF THE COGNITIVE DOMAIN WITH APPROPRIATE VERBS

1. *Knowledge*: The ability to recall learned materials. It can range from the recall of simple facts to complete theories. It represents the lowest level of learning outcome, requiring only that the student recall previously learned information. Statements indicating what students should be able to do could use such terms as:

knows	lists	matches	what
defines	names	outlines	where
identifies	reproduces	states	when
describes	selects	recognizes	who
labels	recalls	memorizes	

2. *Comprehension*: The ability to grasp the meaning of material learned. The student may show understanding of the material by translating it from one form to another by conveying meaning, or by making summary statements about it. Action terms used as indicators of student achievement include the following:

understands	interprets	translates	shows
estimates	justifies	converts	distinguishes
defers	explains	generalizes	extends
infers	predicts	paraphrases	gives examples
rewrites	summarizes	demonstrates	why
how	condenses		

3. *Application*: The ability to use learned materials in new and concrete situations. The student is required to apply rules, concepts, principles, laws, or theories. Indicators of student achievement would include the following:

applies	organizes	solves	relates
constructs	changes	computes	produces
demonstrates	discovers	manipulates	uses
modifies	operates	predicts	

4. *Analysis*: The ability to break material down into its component parts so that its organizational structure may be understood. The student demonstrates attainment of objectives through the ability to identify parts, show relationships, and recognize organizational principles. Objectives written at this level would use words such as the following:

recognizes	categorizes	differentiates	relates
diagrams	distinguishes between	classifies	outlines
illustrates	discriminates	analyzes	

*Adapted from Gronlund (1985)

Appendix A (cont'd)

5. *Synthesis*: The ability to put parts together to form a new whole. The student demonstrates an ability to devise a new plan of operation, or to produce a set of abstract relations. The emphasis is on creativity, and appropriate verbs would include the following:

proposes	integrates	creates
categorizes	combines	compiles
devises	rearranges	reconstructs
composes	formulates	develops

6. *Evaluation*: The ability to judge the value of material. The student might be required to judge the value of a statement, a piece of prose, a poem, an advertisement, or a research report. To indicate that the objective is achieved, the student may be required to:

judge	appraise	conclude
contrast	criticize	justify
interpret	support	

The use of appropriate verbs can be further illustrated as follows:

TOPIC: GRAPHS

Objective:	Students will demonstrate an understanding of graphs.
Knowledge:	Students can identify the information to be graphed.
Comprehension:	Students can interpret information given in a graph.
Application:	Students can use information to construct a graph.
Analysis:	Students can analyze information given in a graph.
Synthesis:	Students can make summary statements or draw conclusions based on information given in the graph.
Evaluation:	Students can judge the adequacy with which conclusions are supported by available data.

APPENDIX B: LEVELS OF THE AFFECTIVE DOMAIN WITH APPROPRIATE VERBS*

Affective Objective: The student will demonstrate an interest in Science.

- Specific Indicators:*
1. The student participates in extra-curricular activities such as science clubs and science fairs.
 2. The student does extra readings in areas related to science.
 3. The student watches and discusses television programs related to science.

Five levels: examples and appropriate verbs.

Receiving: A willingness to attend to specific stimuli.

to accept	to listen	to choose
to select	to ask	to attend

Responding: A willingness to participate,

to answer	to select	to write
to conform	to tell	to follow

Valuing: The worth or value a student attaches to a specific object, behavior, or phenomenon.

to choose	to invite	to share
to support	to join	to appreciate

Organization: Concerned with bringing together different values, resolving conflicts between them, and beginning the construction of an internally consistent value system.

to formulate	to relate	to defend
to abstract	to define	to put in order

Characterization: Deals with patterns of adjustment and is characterized by pervasiveness, consistency, and predictable behavior.

to discriminate	to complete	to behave
to practice	to serve	to verify

*Adapted from Gronlund (1985)

APPENDIX C: LEVELS OF THE PSYCHOMOTOR DOMAIN WITH APPROPRIATE VERBS*

Perception: The use of sense organs to obtain cues that guide motor activity.

Objective: The student can relate music to a particular dance step.

Verbs:	chooses	describes	detects
	separates	identifies	differentiates

Set: Readiness to take a particular type of action.

Objective: The student demonstrates proper bodily stance for batting a ball.

Verbs:	begins	displays	explains
	moves	proceeds	reacts
	responds	shows	starts

Guided Response: Deals with the early stage in learning a complex skill. It includes imitation as well as trial and error.

Objective: The student can perform a golf swing as demonstrated.

Verbs:	assembles	builds	constructs
	dismantles	dissects	fastens
	fixes	measures	manipulates
	mends	sketches	

Mechanism: Deals with performance acts where the learned responses have become habitual and the movements can be performed with some confidence and proficiency.

Objective: The student writes smoothly and legibly.

Verbs: The same list as for guided response.

*Adapted from Gronlund (1985)

Appendix C: (cont'd)

Complex Overt Response: Skillfully performs a motor act that requires complex movement patterns.

Objective: The student can demonstrate correct form in swimming backstroke.

Verbs: The same list as for guided response.

Adaption: Deals with skills which are so well developed that the individual can modify movement patterns to fit special requirements or to meet a problem situation.

Objective: The student modifies his swimming stroke to fit the roughness of the water.

Verbs:

adapts	alters	changes
varies	rearranges	reorganizes
revises		

Origination: Deals with creating new movement patterns to fit a particular situation or specific pattern.

Objective: The student creates a new dance step.

Verbs:

arranges	combines	composes
constructs	creates	designs
originates		

APPENDIX D: A TABLE OF SPECIFICATIONS

CONTENTS		SKILL LEVEL COMPLEXITY							
Chapter/Topic	Sub-Topic	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation	Percent	
Student Politice pp. 1-18	1.1. Who Cares?	3	-	-	-	-	-	3	
	1.2. Why Laws?	1	-	-	-	-	-	1	
	1.3. Student Government	4	1	-	5	-	-	10	
Values, Opinions, and Political Behaviour, pp. 18-40	2.1. What are your Values?	-	-	-	-	-	8	8	
	2.2. When can Gov't Take Away our civil Rights?	5	-	-	-	-	-	5	
	2.3. How do Canadians Express Their Values?	4	6	-	-	4	-	14	
	3.1. What are the Functions of Government?	-	6	-	-	-	-	6	
Governments in Canada pp.41-64	3.2. What are the Different Levels of Gov't in Canada?	-	-	5	-	4	6	15	
	3.3. Reasons for Canada's Federal System of Government in 1867.	6	-	-	-	-	-	6	
	3.4. The Federal Gov't.	5	-	-	-	4	-	9	
	3.5. The Provincial Gov't	-	7	-	-	-	-	7	
	3.6. Judicial Function	-	-	6	-	-	-	6	
	3.7. Municipal Gov't.	-	-	-	-	-	5	5	
	3.8. Fed-Prov. Relations	5	-	-	-	-	-	5	
	TOTALS	33	20	11	5	12	19	100	

APPENDIX E: A GENERAL EVALUATION PLAN

1. *STUDENT POLITICS*

ITEM FORMAT

The student will

- | | | |
|------|--|------------------------------------|
| 1.1 | Identify the appropriate definitions of government, politician, and candidates. | matching |
| 1.2 | Identify the appropriate definition of law. | matching |
| 1.31 | Identify the appropriate definitions of citizenship, matching democracy, nomination and acclamation | matching |
| 1.32 | Distinguish between majority and plurality. | m. choice |
| 1.33 | Given a case study of a student council, extrapolate the elements which made it democratic/undemocratic. | extended-response
or discussion |

2. *VALUES, OPINIONS AND POLITICAL BEHAVIOUR*

The student will

- | | | |
|------|---|--|
| 2.1 | React to an issue and explain what he would do with a supporting rationale. | opinion/simulation
extended-response |
| 2.2 | Know when the government may use the War Measures Act. | debate/values
extended-response
clarification/exercise
supply |
| 2.31 | Provide an appropriate definition of compromise and lobbies. | |
| 2.33 | Research an incident of unjustified police action in a civil disobedience situation. | research paper
and oral report |
| 2.34 | Describe the differences in policies between the Progressive Conservative and Liberal parties on a given issue. | extended-response |

Appendix E (cont'd)

- | | | |
|------|---|--|
| 2.35 | Through use of Gallup Poll results, determine preferences of Canadians by party, by issue, and by region. | supply |
| 2.36 | Apply the poll to his/her neighbourhood and describe results. | extended-response/
survey/discussion/
graphing |
| 2.37 | Suggest possible reasons for patterns evidenced on the poll results. | matching |

3. *GOVERNMENTS IN CANADA*

The student will

- | | | |
|------|---|-------------------------------------|
| 3.1 | Given certain statements about selected government functions, determine whether the function is legislative, executive, or judicial. | matching |
| 3.21 | Given statements by politicians on federalism, will take and defend his/her reaction to the position(s). | debate/discussion
reaction paper |
| 3.22 | Given problem statements, determine which provincial government department should be contacted for action. | matching |
| 3.23 | Given a local problem situation, suggest a plan of action for resolving the issue. | extended-response |
| 3.3 | Recall reasons why a federal system of government was established in 1867. | extended-response |
| 3.41 | Pretend s/he is a news reporter, and formulate questions s/he would ask a senator about the role of the senate in the parliamentary system. | stimulation
extended-response |
| 3.42 | Describe the duties of the Governor-General. | role-playing
extended-response |

Appendix E (cont'd)

- | | | |
|------|---|--------------------------|
| 3.43 | Write a letter to his/her MP to get his/her opinions on an issue for subsequent analysis in class. | letter/discussion |
| 3.5 | Describe the difference between the responsibilities of the Federal and Provincial Government in some area (eg., education, law, health, etc.). | extended-response |
| 3.6 | Given an offense, determine which court of law will hear the case. | supply |
| 3.7 | Take a stand on a local controversial issue. | debate/extended-response |
| 3.8 | Outline some of the contentious issues between essay the Federal Government and the Government. | extended-response |

APPENDIX F-1: CHECKLIST

	Rarely	Sometimes	Usually	Comments
<i>Reading Strategies</i> Expects the text to make sense Searches for meaning Makes predictions Self-corrects Processes chunks of language <i>When in Difficulty:</i> Reads on to end of sentence Re-reads sentence Guesses using initial letter as cue <i>After Reading:</i> Can re-tell story in own words Able to discuss characters Recalls main ideas Understands developing sequence Can describe setting				

APPENDIX F-2: ANECDOTAL RECORD

Name: _____

Date: _____

Grade: _____

Observed Behavior	Comments	Program Plans

APPENDIX F-3: RATING SCALE

MUSIC	
Evaluation of Recorder Playing	
Name: _____ Grade: _____ Date: _____	
<i>RHYTHM</i> Correct rhythm:	_____
Consistent tempo	_____
<i>FINGERING AND NOTES</i> Correct fingering and notation	_____
<i>TONGUEING AND BREATHING</i> Correct phrasing and articulation	_____
<i>POSTURE</i> Stance and holding of recorder	_____
<i>OVERALL PERFORMANCE</i>	_____
KEY: 2 - good (but not necessarily perfect) 1 - satisfactory 0 - unsatisfactory	

APPENDIX G: PROJECT ASSIGNMENT SHEET

Topic Choice(s):

- (a) Assign a specific topic.
- (b) Provide a class of topics from which students can make a selection.
- (c) Allow students to propose their own topics from within a general subject area.

Objectives:

Provide a limited list of primary/specific objectives concerning the learning expectations for students. These objectives should be stated in terminology which is familiar to students.

Task Description:

Include in the task description:

- (a) what students may/may not do;
- (b) how they might do it;
- (c) where work will be completed, and
- (d) what forms are acceptable for a completed project.

Relevant Dates:

(a) Preliminary dates:

- (i) provide several dates for teacher/student discussion;
- (ii) provide dates for peer discussion and consultation, and
- (iii) provide dates for progress reporting.

(b) Final submission date

Evaluation Criteria:

Include the specific criteria on which the project will be evaluated. Indicate the weights assigned to each criterion.

APPENDIX H: STUDENT SELF-EVALUATION FORM

Am I developing into an independent reader?

Name: _____

Grade: _____

Date: _____

Choosing Books

- Do I choose at the correct level?
 - Do I choose a variety of books?
 - Do I listen to the suggestions of others?
 - Do I use all available sources?
-

Reading Independently

- Do I enjoy silent reading time?
 - Do I choose to read at other times?
 - Do I read different books for different purposes?
 - Do I know what to do when I don't understand something
 - Do I know what to do when I don't know a word?
-

APPENDIX I: CONSTRUCTED-RESPONSE ITEMS

Extended-response items: The examinee responds to a question with a paragraph or more. These items are often referred to as essay items.

Guidelines for Construction:

1. Base questions on the objectives of the unit of work being tested.
2. Questions requiring higher-order mental processes can not be a repetition of situations or problems used during the instruction.
3. Wording of the question should be specific and clear. Examinees should know exactly what they are to do.
4. Note on the examination the number of marks that each essay item is worth. Marks allotted should reflect the difficulty of the item and the time required to answer it.
5. To obtain adequate reliability, it is necessary to prepare a scoring guide beforehand.
6. Examinees should be made aware beforehand of the criteria against which their essays will be evaluated.

Restricted-response items: The examinee responds to a question with a word, sentence, a brief paragraph or a list of statements. This includes the shorter essay items or completion objectives.

Guidelines for Construction: In addition to those stated above which may apply to many restricted-response items, the following should be noted for completion items.

1. Items should require a single-word answer rather than a phrase. Synonyms should be keyed as correct.
2. Only key words should be left blank.

3. The blank should appear at, or very near, the end of the statement. This ensures that the students will have all necessary information required to provide the answer. It also facilitates scoring.
4. No cues to the answer should be provided. If the word *an* appears before a blank, for example, it cues the student that the answer begins with a vowel. Therefore a/an should be used.
5. Where units are required, they should be put in.
6. Statements should not be lifted directly from the text, since doing this encourages rote memorization.
7. One element only should be omitted from a statement.
8. All blanks should be of uniform length. A single blank should always be used even when the answer is two words such as New York.
9. Statements should be definite such that they cannot have more than one correct answer.

APPENDIX J-1: SELECTED-RESPONSE ITEMS

True-False

Description: The examinee indicates whether a statement is true or false.

General: These items are best suited for testing knowledge but are very difficult to write for some subjects such as the social sciences. Tests comprised of true-false items do not produce reliable results, but reliability can be improved by including many items. Another limitation is that it has no diagnostic function. They are best when only two choices exist such as right or left, distinguishing between cause and effect and distinguishing fact from opinion.

Guidelines for construction:

1. Avoid absolutes like *all* and *never*. Statements containing such words are usually false, and the use of qualifiers cues the examinee to answer true.
2. Key approximately one-half the items true.
3. Avoid making statements false by insertion of the word *not*. If a negative has to be used, the negative should be highlighted.
4. Avoid double negatives.
5. Make sure that there are no cues for students (such as true statements always being longer than false ones).
6. Statements should be short, and embody one idea.
7. Avoid taking statements directly from a text.
8. If opinion is used, it should be attributed to the source.
9. Statements must be clearly true or clearly false.
10. Statements must not include indefinite terms, degrees, or amounts.
11. To make false statements sound true, phrases which give them a *ring of truth* may be used.

APPENDIX J-2: SELECTED-RESPONSE ITEMS

Matching

Description: This format contains two parallel columns with each word, phrase, number or symbol in one column being matched with a word or phrase in the other column.

General: Matching items are limited to measuring knowledge which is based on simple associations. They are appropriate where an objective requires "the ability to identify the relationships between two things". Thus, they are restricted to relatively small areas of achievement. Matching items are thought to be easy to write, but they are often made difficult by the requirement that each response be a plausible answer for each premise. The advantage of matching items is that they can measure knowledge of much content in a short time.

Guidelines for construction:

1. The lists should have more responses than premises.
2. Responses should be placed in some logical order (alphabetical, for example).
3. Each alternative must be a plausible solution to each premise.
4. Instructions should be clearly stated. If responses may be used more than once, it should be indicated.
5. All items should be on one page.
6. Responses should be as concise as possible. Verbatim statements from the text should not be used.
7. Difficulty level may be controlled by the discriminations the examinee is expected to make. Generally speaking, the broader the terms in the premises list, the easier the item is.
8. The responses and premises should contain no cues such as plural forms.
9. There should be 8-10 matching items in one set.
10. Both columns should be labelled.

APPENDIX J-3: SELECTED-RESPONSE ITEMS

Multiple-choice

Description: The student indicates which one of three or more choices is correct or most correct.

General: Multiple-choice items can be used to measure more than recall of knowledge. (See earlier discussion in this section.) These items are adaptable to most subjects. If they are well-written, they are easily understood by students, and they are easy to score.

Guidelines for construction:

1. The stem should be a complete statement about the problem to be solved.
2. The stem should contain no unnecessary or irrelevant information.
3. The stem should be stated in clear and precise language.
4. The stem should contain as much of the item as possible such that information will not have to be repeated in each alternative.
5. Negative statements should be avoided if possible. If they cannot be avoided, the negative should be placed near the end of the stem and highlighted.
6. All distractors should be consistent with the grammatical and syntactical construction of the stem.
7. Distractors should be approximately equal in length.
8. If an opposite is used as one distractor, the other two distractors should be opposites as well.
9. All distractors should be plausible answers to the problem presented in the stem.
10. There should be no grammatical cues to the right answer.
11. One alternative should be clearly right and the distractors clearly wrong except in the case of the best answer type.
12. Avoid using camouflage distractors (none of the above, all of the above)

Appendix J-3 (cont'd)

13. Modifiers such as sometimes and usually should be avoided.
14. If the correct answer has important-sounding words, distractors should contain such words as well.
15. Choices should be listed vertically and on separate lines.
16. Letters should be in front of choices.
17. Statements should not be lifted directly from the text.
18. Items should contain no tricks.
19. Items that test higher-order mental processes must have an element of novelty, but situations presented in the item should be similar to those used in class.
20. Use the *best* answer-type only when there is no correct answer.

APPENDIX K: LOW ORDER AND HIGH ORDER SELECTED-RESPONSE ITEMS

Lower-order: Lower-order instructional objectives involve knowledge, comprehension application, and analysis.

1. *Knowledge* items require students to recall exactly what was learned, as in the item below.

Objective: Students will match chemical elements with their symbols.

Match each chemical element found in column II with its correct symbol found in column I. Place the number which corresponds with the correct answer in the blank provided to the right of column II.

<u>I</u>	<u>II</u>	
1. Ag	1. Copper	—
2. Au	2. Gold	—
3. Cu	3. Lead	—
4. Hg	4. Mercury	—
5. K	5. Sodium	—
6. Na	6. Silver	—
7. Pb	7. Tin	—
8. Si		
9. Sn		

-
2. *Comprehension* test items may require students to demonstrate understanding of facts and principles. Understanding is demonstrated when descriptions of these facts and principles are presented in a way which is different from the way they were presented in class. Descriptions may be paraphrased or written in different language, for example. They may also require students to interpret verbal material, to interpret charts and graphs, and so on. An example of a comprehension item appears below.

Objective: Students will interpret information which is plotted on a line graph.

Which month has the lowest average precipitation? (Graph not provided)

- (A) January
(B) February
(C) May
(D) December
-

Appendix K (cont'd)

3. *Application* items may require students to construct charts and graphs, to solve mathematical problems, apply principles to new situations, and so on. An application item appears below.
-

Objective: Students will apply the principle of supply and demand to situations.

If the demand for a commodity or service is relatively constant, a decrease in its supply will result in an increase in its market value.

Directions: Read each of the following statements. For any statement which agrees with the above principle, mark an X in the blank provided.

1. Since world war II, there has been an upward trend in the price of stocks. _____
 2. Fresh fruit and vegetables cost more when not in season. _____
 3. The cost of medicine is higher now than it was ten years ago. _____
-

4. *Analysis* items may require students to recognize unstated assumptions, recognize logical fallacies in reasoning, distinguish between facts and influences, analyze the organizational structure of a work, and so on. It should be noted that analysis items cannot be answered through information which is provided in the test item itself. Students must actually do some analysis; otherwise, a different cognitive ability is being tested. An analysis item appears below.
-

Objective: Students will analyze a poem to determine its tone.

What is the tone of the poem?(poem not provided)

- (A) jaunty and bantering
 - (B) sarcastic
 - (C) silly
 - (D) sorrowful
-

Appendix K (cont'd)

Higher Order: Selected-response items are not suitable for measuring students' ability to engage in such cognitive processes as synthesis and evaluation. These abilities are better measured through the use of constructed-response items. It is possible, however, to write selected-response items which test students' understanding of concepts, principles, and methods and procedures. The key to writing items to test understanding is the addition of elements which would be new to the student. It should be noted, however, that although situations presented in the item should be different from those used in teaching a course, they should not be too far removed from the examples which a student would have encountered in class. Several illustrative items are presented below.

1. Understanding of principles (if-then relationships):

Objective: Given a previously *unused* example of the application of a principle, select the correct name from a set of alternatives.

What causes soup to become thicker each time it is reheated?

- (A) condensation
 - (B) evaporation
 - (C) radiation
 - (D) refraction
-

2. Understanding of Methods and Procedures:

Objective: Given a cue, students shall select from a list of alternatives, the best of several possible explanations of a method or procedure.

Why is adequate lighting necessary in a balanced aquarium?

- (A) Fish need light to see their food.
 - (B) Fish take in oxygen in the dark.
 - (C) Plants expel carbon dioxide in the dark.
 - (D) Plants grow too slowly in the dark
-

Appendix K (cont'd)

3. Understanding of Concepts

Objective: Given the name, select a previously *unused* example of a concept from a set of alternatives.

Which of the following is a chair?

- (a) lounge
 - (b) loveseat
 - (c) rocker
 - (d) stool
-

It should be noted that these examples are meant to be illustrative only. The objective to be tested in assessing understanding of principles could be stated in reverse, as in the example below.

Objective: Given the name, select from a set of alternatives a previously *unused* example of an application of a principle.

The principle of capillary action helps explain how fluids do which of the following?

8.
 - (a) enter solutions of lower concentration
 - (b) escape through small openings
 - (c) pass through semipermeable membranes
 - (d) rise in tubes
-

In testing understanding of concepts, the distracters should be near-examples of the correct response. Each of the distracters should be missing only one of the critical attributes of the concept. The critical attributes of a chair are:

- (1) It seats only one person.
- (2) It has support for the back.
- (3) It seats a person in an upright position.

A lounge does not seat a person in an upright position, a loveseat seats more than one person, and a stool has no support for the back. Thus, while they are near-examples of chairs, they are not examples. (For a more detailed discussion of concepts, see Roid and Haladyna, 1982).

APPENDIX L*: GRADE RETENTION SCALE

NAME _____ **DATE** _____

D.O.B. _____ **GRADE SCHOOL** _____

INSTRUCTIONS: Score by checking (✓) each factor. Weigh in favour of non-retention (-1,-2,-3), weigh in favour of retention (+1, +2, +3) or zero (0) if factor is neutral or does not apply. If information is insufficient to respond to a factor, score 0. The closer to -47 a total is, the more certain one can be that a student should be promoted. The closer to +29 it is, the more certain one can be that retention is the best solution. The closer to 0 it is, the less certain the decision.

*CEA Information Note
Canadian Education Association, March, 1989.

		Non-Retention			Retention			
		-3	-2	-1	0	+1	+2	+3
1.	Previous Retentions Score (-3) if student has repeated one of more grades, otherwise score (0)							
2.	Physical Disabilities Score (-3) if present, otherwise score (0)							
3.	Learning Disabilities Score (-3) if present, otherwise score (0)							
4.	Physical Size Score (-1) if stature is large for age; score (+1) if small for age, otherwise score (0)							
5.	Academic Potential Score (-2) when slow learning attributes are present; score (-1) if chronically underachieves Score (+1) if learning occurs after prolonged practice; score (+2) if attributes of slow learning are temporary							
6.	Psychosocial Maturity Score (-3) if mature; score (-1) if mature and peer difficulties; score (+1) if immature and age 9 or over; score (+2) if immature and age 7.6 to 8.6; score (+3) if immature and age 6 to 7.5							
Subtotal page 1								

7. **Neurological Maturity**

Score (-2) if no indicators of immaturity are present

Score (0) if one or more indicators of immaturity are present and age 12 or over

Score (+1) if one or more indicators of immaturity are present and age 10 to 11

Score (+2) if one or more indicators of immaturity are present and age 8 to 9

Score (+3) if one or more indicators of immaturity are present and age 6 to 7

Non-Retention			Retention			
-3	-2	-1	0	+1	+2	+3

--	--	--	--	--	--	--	--

8. **Self-Concept**

Score (-3) if low self-concept and age 9 or over; score (-2) if good self-concept; score (0) if low self-concept is result of low achievement and age greater than 11; score (+2) if low self-concept is result of low achievement and age 10 to 11; score (+3) if low self-concept is result of low achievement and age 9 or less

--	--	--	--	--	--	--	--

9. **Ability to Function Independently**

Score (-3) if student functions independently; score (+1) if student cannot function independently and is age 11 or over; score (+2) if student cannot function independently and is age 10 or younger

--	--	--	--	--	--	--	--

10. **Grade Placement**

Score (-3) if student is in grade 5 or beyond; score (0) if student is in grade 4; score (+3) if student is in grades 1, 2 or 3

--	--	--	--	--	--	--	--

11. **Chronic Absenteeism**

Score (0) if not a problem: score (+3) if this is a problem

--	--	--	--	--	--	--	--

12. **Acute Behaviour Problems**

Score (-3) if present: score (0) otherwise

Subtotal page 2

13. **Basic Academic Skill Deficiencies**
Score (0) if not present; score (+3) if present

Non-Retention				Retention			
-3	-2	-1	0	+1	+2	+3	

--	--	--	--	--	--	--	--

14. **Peer Pressure**
Score (-3) if student is a) susceptible to peer pressure or b) has close friendships with classmates or c) lives close to agemates/classmates, otherwise score (0)

--	--	--	--	--	--	--	--

15. **Student's Attitude Toward Retention**
Score (-3) if strongly opposed; score (0) if neutral; score (+3) if student agrees

--	--	--	--	--	--	--	--

16. **Emotional Development**
Score (-3) if student experiences chronic emotional stress, otherwise score (0)

--	--	--	--	--	--	--	--

17. **Geographic Moves Within Past Two School Years**
Score (-2) if no moves; score (0) if one or two moves; score (+2) if more than two moves

--	--	--	--	--	--	--	--

18. **Non-English Language Student**
Score (0) if English speaking; score (+2) if English is second language

--	--	--	--	--	--	--	--

19. **Parents' Attitude Toward Retention**
Score (-3) if strongly opposed; score (+3) if strongly agree

--	--	--	--	--	--	--	--

20. **Age of Siblings**
Score (-3) if sibling(s) are same age (twin) or one to two years younger; score (0) if no siblings are older

--	--	--	--	--	--	--	--

Subtotal page 3

21. Availability of Personnel

Score (-2) if student must repeat with same teacher in absence of teacher-learner rapport; score (+2) if student can be retained with alternative teacher, or score (+2) in presence of teacher-learner rapport; or score (+2) if school transfer and "fresh start" is likely to occur

Non-Retention				Retention			
-3	-2	-1	0	+1	+2	+3	

Subtotal page 4

Subtotal page 1

Subtotal page 2

Subtotal page 3

Total

Add

Grand Totals

(-)

+ADD+

(+)

Achieved Score (+-)

Principal's Signature

Teacher's Signature

Other